APR 20 1936 C-27

AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

J. F. RITT
COLUMBIA UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY	MARSTON MORSE	ALONZO CHURCH
HARRY BATEMAN	J. R. KLINE	L. R. FORD
W. A. MANNING	E. P. LANE	OSCAR ZARISKI
HARRY LEVY	AUREL WINTNER	G. C. EVANS

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LVIII, Number 2

APRIL, 1936

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

Waring theorems of new type. By L. E. DICKSON, . . . 241
Fundamental systems of units in normal fields. By MARIE J. WEISS, . 249
A calculus of sequences. By MORGAN WARD, ... 255
On the transformation and convergence of continued fractions. By WALTER LEIGHTON and H. S. WALL, ... 267
The form $2wx + xy + yz + zu + ux$. By E. T. BELL, ... 282
Reducible exceptional curves. By PATRICK DU VAL, ... 285
Theorems concerning a basis for all solvable groups, with a list of basis groups of degree < 64. By A. C. LUNN and J. K. SENIOR, 290
Arc-preserving transformations. By G. T. WHYBURN, ... 305
An analytic characterization of surfaces of finite Lebesgue area. Part II. By CHARLES B. MORREY, JR., ... 313
Some separation properties of the plane. By R. E. BASYE, . . . 323
The analysis of certain curves by means of derived local separating points. By G. E. SCHWEIGERT, ... 329
On the independence of Hilbert and Ackermann's postulates for the calculus of propositional functions. By J. C. C. MCKINSEY, 336
An unsolvable problem of elementary number theory. By ALONZO CHURCH, ... 345
Asymptotic solutions of certain ordinary differential equations associated with multiple roots of the characteristic equation. By HUGH L. TURRITTIN, ... 364
Biharmonic functions and certain generalizations. By EDWARD KASNER, 377
The problem of Bolza in the calculus of variations in parametric form. By MAGNUS R. HESTENES, ... 391
On expansions in terms of a certain general class of functions. By W. T. MARTIN, ... 407
On the boundary of the range of values of $\zeta(s)$. By RICHARD KERSHNER and AUREL WINTNER, ... 421
The sampling theory of systems of variances, covariances and intraclass covariances. By S. S. WILKS, ... 426
Totally disconnected locally compact rings. By N. JACOBSON, . . . 433
On singular Fourier-Stieltjes transforms. By RICHARD KERSHNER, . 450

THE AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

94
1
a 461

WARING THEOREMS OF NEW TYPE.

By L. E. Dickson.

---

1. Denote $ax^3 + by^3$ by (a, b) and similarly for several variables, which we restrict to integral values $\geqq 0$. We shall prove

Theorem 1. *Every integer $\geqq 0$ is a sum of three values of any of the forms $(1, 1, j)$, $j = 1, 2, 5$; $(1, 2, k)$, $k = 2, \cdots, 6$.*

For $(1, 1, 1)$, this is Waring's theorem on nine cubes.

Theorem 2. *Every integer $\geqq 0$ is a sum of five values of $(1, 2)$. All integers $> 11^{12}$ (19.006868) are sums of five values of $(1, 5)$.*

In Theorem 2 we may take zero as one of the ten cubes.

Call a form universal if it represents all integers $\geqq 0$. Theorem 5 gives 6344 universal forms each a sum of nine products of a cube by a positive integer. Certain of them yield Theorem 1.

Our method is a direct generalization of the classic one for nine cubes.[1] For some of our forms the auxiliary work in Lemma 3 is much simpler than that for nine cubes. The problem is under investigation by different methods.

2. *The four lemmas.*

Lemma 1. *If p is a prime $\equiv 2 \pmod 3$, every integer not divisible by p is congruent to a cube modulo p^n.*

This well known Lemma (*Transactions, loc. cit.*) implies

Lemma 2. *If $p \equiv 2 \pmod 3$, and if l is not divisible by the prime p, every integer not divisible by p is congruent modulo p^n to the product of a cube by l.*

Consider the forms

$$A = x^2 + y^2 + z^2, \quad B = x^2 + y^2 + 2z^2, \quad C = x^2 + y^2 + 5z^2,$$
$$D = x^2 + 2y^2 + 2z^2, \quad E = x^2 + 2y^2 + 3z^2, \quad F = x^2 + 2y^2 + 4z^2,$$
$$G = x^2 + 2y^2 + 5z^2, \quad H = x^2 + 2y^2 + 6z^2.$$

---

[1] Dickson, *Transactions of the American Mathematical Society*, vol. 30 (1928), pp. 1-18. The discussion on pp. 7-13 becomes simpler by our new Lemma 4 since we now obtain the final C at once without reducing a preliminary C. As on p. 14, we readily extend our Lemma 3 to further values of P.

RJ

LEMMA 3. *Let P and e be given integers $\geqq 0$. If $P \equiv 5 \pmod{16}$, every integer $\geqq P^e \cdot 11^3$ is represented by $P^e\gamma^3 + 6f$, where $\gamma \geqq 0$ and f is B, E, or F. If $P \equiv 5 \pmod{48}$, every integer $\geqq P^e \cdot 22^3$ is represented by $P^e\gamma^3 + 6f$, where $\gamma \geqq 0$ and f is A, D, or H. If $P \equiv 11 \pmod{48}$, every integer $\geqq P^e \cdot 23^3$ is represented by $P^e\gamma^3 + 6f$, $\gamma \geqq 0$, $f = A$ or C. If $P \equiv 11 \pmod{50}$, every integer $\geqq P^e \cdot 11^3$ is represented by $P^e\gamma^3 + 6G$ with $\gamma \geqq 0$.*

The proof is simplest for B, E, F, or G. Elsewhere [2] I proved that E represents exclusively all positive integers not of the form $4^k(16r + 10)$. Hence E represents every positive integer congruent modulo 16 to

$$1, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14, 15, \tag{1}$$

or 2, which we omit. Let (γ, μ) denote the least positive residue of $\gamma^3 + 6\mu$ modulo 96. The 72 distinct numbers (γ, μ), with $0 \leqq \gamma \leqq 5$ and μ ranging over (1), together with

$$\begin{gathered}(6,4) = 48,\ (6,6) = 60,\ (6,12) = 0,\ (6,14) = 12,\ (7,1) = 61,\\ (7,7) = 1,\ (7,9) = 13,\ (7,15) = 49,\ (8,4) = 56,\ (8,6) = 68,\\ (8,12) = 8,\ (8,14) = 20,\ (9,3) = 75,\ (9,5) = 87,\ (9,11) = 27,\\ (9,13) = 39,\ (10,4) = 64,\ (10,6) = 76,\ (10,12) = 16,\ (10,14) = 28,\\ (11,1) = 89,\ (11,7) = 29,\ (11,9) = 41,\ (11,15) = 77,\end{gathered}$$

are found to be a rearrangement of $0, 1, \cdots, 95$. Since the product of the numbers (1) by 5 (and hence by P) are congruent modulo 16 to the same 12 numbers permuted, we conclude that, if n is any integer,

$$n \equiv P^e\gamma^3 + 6\mu \pmod{96} \tag{2}$$

has integral solutions with $0 \leqq \gamma \leqq 11$, μ in (1). Thus there is an integer q for which

$$n = P^e\gamma^3 + 6\mu + 96q = P^e\gamma^3 + 6m, \qquad m = \mu + 16q. \tag{3}$$

Let $n \geqq P^e \cdot 11^3$. Then $n \geqq P^e\gamma^3$, $m \geqq 0$, whence m is represented by E.

Proof for B. The form B represents exclusively all positive integers not of the form $4^k(16r + 14)$. Hence B represents all congruent modulo 16 to

$$1, 2, 3, 4, 5, 7, 9, 10, 11, 12, 13, 15, \tag{4}$$

or 6. The 72 residues (γ, μ), with $0 \leqq \gamma \leqq 5$ and μ in (4), together with

[2] *Bulletin of the American Mathematical Society*, vol. 33 (1927), pp. 63-70.

$(6,2)=36,\ (6,4)=48,\ (6,9)=78,\ (6,10)=84,\ (6,12)=0,$
$(7,5)=85,\ (7,7)=1,\ (7,13)=37,\ (7,15)=49,\ (8,2)=44,$
$(8,4)=56,\ (8,10)=92,\ (8,12)=8,\ (9,1)=63,\ (9,3)=75,$
$(9,9)=15,\ (9,11)=27,\ (10,2)=52,\ (10,4)=64,\ (10,10)=4,$
$(10,12)=16,\ (11,5)=17,\ (11,7)=29,\ (11,13)=65,\ (11,15)=77,$

form a rearrangement of $0, 1, \cdots, 95$. The products of (4) by 5 are congruent modulo 16 to the same 12 numbers rearranged. Hence (2) is solvable with $0 \leqq \gamma \leqq 11$, μ in (4). We have (3) with m represented by B. This proof applies unchanged also to F.

Proof for G. By *Bulletin, loc. cit.*, G represents exclusively all positive integers $\neq 25^k(25r \pm 10)$ and hence represents all $\equiv \mu \pmod{50}$, $\mu = 1, \cdots, 9,\ 11, \cdots, 14,\ 16, \cdots, 24,\ 26, \cdots, 34,\ 36, \cdots, 39,\ 41, \cdots, 49$. Since $11 \times 5 \equiv 5$, multiplication yields $11j \equiv j \pmod{50}$ for $j = 5, 20, 30, 45$, i. e., for all the values of μ which are multiples of 5. When x ranges over the integers prime to 5 and < 50, the same is true of the residues of $11x$ modulo 50. Hence μ and 11μ take the same residues modulo 50. For $0 \leqq \gamma \leqq 11$, the residues modulo 300 of $\gamma^3 + 6\mu$ are found to be $0, 1, \cdots, 299$. Multiplication by $P \equiv 11 \pmod{50}$ shows that $n \equiv P^e\gamma^3 + 6\mu \pmod{300}$ is solvable for every n with $0 \leqq \gamma \leqq 11$.

Proof for H. The form H represents [3] exclusively all positive integers not of the form $4^k(8s+5)$. Hence H represents all $\equiv$

(5) $\qquad 2,\ 3,\ 6,\ 7,\ 8,\ 10,\ 11,\ 12,\ 14,\ 15,$

or 1 or 9 (mod 16). Employ $0 \leqq \gamma \leqq 5$, all μ in (5); $\mu = 7, 11$, $\gamma = 7, 11, 15$; $\mu = 11, 15$, $\gamma = 9, 13, 17$; $\mu = 8$, $\gamma = 7, 9, 10, 11, 18$; $\mu = 12$, $\gamma = 6, \cdots, 9, 11, 14, 22$. The resulting residues (γ, μ) form a permutation of the numbers $0, \cdots, 95$ except $6 + 8r$ $(r = 0, \cdots, 11)$. The latter are the values of (γ, μ) for $\mu = 1, 9$, $\gamma = 0, 2, 4, 6, 8, 10$. Hence, for $e = 0$, (2) is solvable with $0 \leqq \gamma \leqq 22$ and $\mu = 1, 9$ or one of (5). In $5(6 + 8r) = 6 + 8\rho$, $\rho = 3 + 5r$ ranges with r over a complete set of residues modulo 12. Hence multiplication by 5 of the 12 numbers $6 + 8r$ or of the 10 numbers (5) merely permutes them modulo 96. Evidently the same holds when 5 is replaced by P. The final step of the proof for H is the same as for A (*Transactions, loc. cit.*). That proof for A applies unchanged to D since A and D each represent exclusively all positive integers $\neq 4^k(8s+7)$.

[3] B. W. Jones, *Chicago Dissertation* (unpublished), 1928, p. 72.

Proof for C. The form C represents [4] exclusively all positive integers $\neq 4^k(8s+3)$. Hence C represents all $\equiv 1, 4, 9$ or

$$2,\ 5,\ 6,\ 7,\ 8,\ 10,\ 13,\ 14,\ 15, \tag{6}$$

modulo 16. Employ $0 \leqq \gamma \leqq 5$, all μ in (6); $\mu = 5, 7, 8, 13, 15$, $\gamma = 6, 8, 10$; $\mu = 6, 14$, $\gamma = 9, 13, 17, 21$; $\mu = 2, 7, 8, 10$, $\gamma = 7, 11$; $\mu = 7, 15$, $\gamma = 9$; $\mu = 2$, $\gamma = 15, 19, 23$; $\mu = 8$, $\gamma = 14, 18, 22$. The resulting residues (γ, μ) give $0, \cdots, 95$ except $0, 32, 64$. The proof for C is now completed as for A in (*Transactions, loc. cit.*, p. 9), since the products of (6) by 11 are congruent modulo 16 to the same nine numbers permuted.

LEMMA 4. *Given R and the positive numbers s and t such that*

$$s \geqq t + t/9^3, \qquad t \leqq R \leqq s, \tag{7}$$

we can find an integer $i \geqq 0$ satisfying

$$R \leqq s - ti^3 < R + K, \qquad K = 3[t(s-t)^2]^{1/3}. \tag{8}$$

If $s < R + K$, (8) holds for $i = 0$. Henceforth, let $s \geqq R + K$. Determine a real number r so that $s - tr^3 = R$. By (7_1),

$$tr^3 \geqq K \geqq t/27, \quad r \geqq 1/3.$$

We may write $r = i + f$, where $0 \leqq f < 1$ and i is an integer $\geqq 0$. Since $i \leqq r$, $-r^3 \leqq -i^3$ and $R \leqq s - ti^3$, as desired in (8). Next,

$$s - ti^3 - R = s - t(r-f)^3 - s + tr^3 = tw,$$
$$w = r^3 - (r-f)^3 = 3r^2f - f^2(3r-f) < 3r^2f < 3r^2,$$

since $3r - f > 0$. By $t \leqq R$, $tr^3 \leqq s - t$, whence

$$3tr^2 \leqq K, \ s - ti^3 - R = tw < 3tr^2 \leqq K.$$

3. *General theory.* When $C > 0$ and $p > 1$, Cp^{3x} increases indefinitely with x. To any $s - t > 0$ therefore corresponds an integer $n \geqq 0$ for which

$$Cp^{3n} \leqq s - t < Cp^{3(n+1)}. \tag{9}$$

THEOREM 3. *Let s, t, l, h_1, h_2, h_3 be given positive integers for which $s \geqq t + t/9^3$. Choose a prime p, $p \equiv 2 \pmod 3$, which divides neither t nor l, such that, as a composite of the cases in Lemma 3 with $P = p$, every integer $\geqq p^e g^3$ is represented by $p^e\gamma^3 + 6f$ for every integer $e \geqq 0$, where $\gamma \geqq 0$ and*

[4] Dickson, *Annals of Mathematics* (2), vol. 28 (1927), pp. 340-341.

$$f = h_1x_1^2 + h_2x_2^2 + h_3x_3^2. \tag{10}$$

Restrict e and n to integral values for which

$$e = 0, 1, \text{or } 2, \qquad e + n \equiv 0 \pmod 3. \tag{11}$$

Choose ν and C so that, when $n \geqq \nu$ (see § 4),

$$(C - l - w)p^{2n-e} \geqq g^3,\ Cp^{3n} + 2k \leqq (6 + w)p^{3n},\ Cp^{3n} \geqq t, \tag{12}$$

where

$$w = 2(h_1 + h_2 + h_3),\ k = 3(tC^2)^{1/3}p^{2n+2}. \tag{13}$$

Determine n by (9). *Then if $s \geqq t + Cp^{3\nu}$, there exist nine integers a, b, c, y_i, z_i $(i = 1, 2, 3)$, each $\geqq 0$, such that*

$$s = ta^3 + lb^3 + c^3 + \sum_{i=1}^{3} h_i(y_i^3 + z_i^3). \tag{14}$$

The final inequality (9) gives $K < k$ for K in (8).

First, let $Cp^{3n} + 2k \leqq s$. Then Cp^{3n} and $Cp^{3n} + k$ are values of R satisfying (7_2). Hence by (8) there exist integers I and J, each $\geqq 0$, such that

$$Cp^{3n} \leqq s - tI^3 < Cp^{3n} + k,$$
$$Cp^{3n} + k \leqq s - tJ^3 < Cp^{3n} + 2k.$$

Since these intervals contain no common number, $I \neq J$. Hence there exist two distinct integral values of i which satisfy

$$Cp^{3n} \leqq s - ti^3 < Cp^{3n} + 2k, \qquad i \geqq 0. \tag{15}$$

Second, let $Cp^{3n} + 2k > s$. By (9_1), (15) holds when $i = 0$ or 1. Hence in both cases there exist two distinct integers and hence two consecutive integers $j - 1$ and j which are both values of i satisfying (15). At least one of the integers $s - t(j-1)^3$ and $s - tj^3$ is not divisible by p. In fact, their difference is the product of t by $3j^2 - 3j + 1$. If $p = 2$, this is not divisible by p. If $p > 2$,

$$12(3j^2 - 3j + 1) = (6j - 3)^2 + 3$$

is not divisible by p since the reciprocity law gives

$$(-3 \mid p)(p \mid 3) = 1, \quad (p \mid 3) = (2 \mid 3) = -1, \quad (-3 \mid p) = -1.$$

Hence there exists an integer $a \geqq 0$ such that (15) holds when $i = a$, and such that $s - ta^3$ is not divisible by p. By Lemma 2, $s - ta^3 \equiv l\beta^3 \pmod{p^n}$, where β is not divisible by p. Evidently $\beta \equiv b$, where

$$s - ta^3 = lb^3 + p^n M, \quad 0 < b < p^n. \tag{16}$$

By (12_2) (16) and (15) with $i = a$,

$$Cp^{3n} \leqq lb^3 + p^n M < (6+w)p^{3n}, \quad (C-l)p^{3n} < Cp^{3n} - lb^3,$$
$$(C-l)p^{3n} < p^n M < (6+w)p^{3n}.$$

Cancel p^n and write $M = N + wp^{2n}$. Thus

$$(C-l-w)p^{2n} < N < 6p^{2n}, \tag{17}$$

$$s = ta^3 + lb^3 + p^n(N + wp^{2n}). \tag{18}$$

By an assumption in Theorem 3, N is represented by $p^e\gamma^3 + 6f$ if $N \geqq p^e g^3$ and hence by (17_1) if (12_1) holds. By (11), $p^{n+e}\gamma^3$ is the cube of an integer $c \geqq 0$. Thus

$$p^n N = c^3 + 6fp^n, \tag{19}$$

$$s = ta^3 + lb^3 + c^3 + p^n(wp^{2n} + 6f). \tag{20}$$

Since $w = 2\Sigma h_i$,

$$\sum_{i=1}^{3} h_i\{(p^n + x_i)^3 + (p^n - x_i)^3\} = wp^{3n} + 6fp^n. \tag{21}$$

Evidently (20) now gives (14). If $x_i > p^n$ for a certain i, then $f > p^{2n}$ by (10), and $N > 6p^{2n}$ by (19), contrary to (17). Hence each of our nine cubes is $\geqq 0$.

4. *Conditions* (12). For[5] $n \geqq 3$, the minimum of $2n - e$ is 6, and (12_1) holds if

$$C - l - w \geqq (g/p^2)^3. \tag{22}$$

Henceforth let $n \geqq 4$. Then $1/p^{n-2} \leqq 1/p^2$ and (12_2) holds if

$$C + (tC^2)^{1/3}(6/p^2) \leqq 6 + w. \tag{23}$$

For example, consider D. Here $p = 5$, $g = 22$, $w = 10$. First, let $l \geqq 4$. Then $C > 14\frac{2}{3}$ by (22), $C^2 > 215$, $C^{2/3} > 5.99$ and (23) fails. When $l = 3$, $C \geqq 13.68147$ and (23) gives $t \leqq 4$.

In this way we obtain the complete solution of (12) when $n \geqq 4$:

---

[5] Let $n = 3$. We have (23) with p^2 replaced by p. This case is excluded for forms B, D, E, F, G. But for C, it arises only when $l = 2$, $t = 1$; $l = 1$, $t = 1, 2, 3$. For H, there are many sets l, t. Any set with $\nu = 3$ is evidently a set with $\nu = 4$. We obtain universal theorems for all the latter sets. Hence we ignore those special sets which are also sets with $\nu = 3$, in spite of the lowering of the constant $Cp^{3\nu}$.

B, $p=5$.	$l=4$, $t\leqq 3$;	$l=3$, $t\leqq 14$;	$l=2$, $t\leqq 42$;
		$l=1$, $t\leqq 104$.	
C, $p=11$.	$l=5$, $t\leqq 22$;	$l=4$, $t\leqq 200$;	$l=3$, $t\leqq 760$;
		$l=2$, $t\leqq 2038$;	$l=1$, $t\leqq 4533$.
D, $p=5$.	$l=3$, $t\leqq 4$;	$l=2$, $t\leqq 16$;	$l=1$, $t\leqq 42$.
E, $p=5$.	$l=4$, $t=1$;	$l=3$, $t\leqq 7$;	$l=2$, $t\leqq 21$;
		$l=1$, $t\leqq 50$.	
F, $p=5$.	$l=4$, $t=1$;	$l=3$, $t\leqq 6$;	$l=2$, $t\leqq 16$;
		$l=1$, $t\leqq 37$.	
G, $p=11$.	$l=5$, $t\leqq 18$;	$l=4$, $t\leqq 163$;	$l=3$, $t\leqq 612$;
		$l=2$, $t\leqq 1619$;	$l=1$, $t\leqq 3545$.
H, $p=5$.	$l=3$, $t=1$;	$l=2$, $t\leqq 6$;	$l=1$, $t\leqq 15$.

THEOREM 4. *For all these sets l, t such that t is not divisible by p, and for the least C determined by* (22), *conditions* (12) *hold, so that the final statement in Theorem* 3 *is proved.*

5. *Universal theorems.* We employ a lemma whose proof (*Transactions*, p. 4) is very similar to that of our Lemma 4.

LEMMA 5. *Given the positive numbers s and t and a number T for which $0\leqq T\leqq s$, $t\leqq 9^3 s$, we can find an integer $i\geqq 0$ satisfying*

$$T\leqq s-ti^3<T+3(ts^2)^{1/3}. \tag{24}$$

Consider the unfavorable case $G=x^2+2y^2+5z^2$, $p=11$, $l=1$, $t=\tau$, $\tau\leqq 3545$. By (22), the minimum C is 17.00075. The form [6] $J=(11122)$ represents all positive integers $<40{,}000$. For $3s^{2/3}=40{,}000$, Lemma 5 with $T=0$, $t=1$, shows that $s=i^3+J$ if $\log s=6.1874082$. Apply Lemma 5 with $T=0$, $t=5$, and s replaced by σ. Hence if $3(5\sigma^2)^{1/3}=s$, viz., $\log\sigma=8.2159453$, all positive integers $\leqq\sigma$ are represented by $(1_4, 2_2, 5)$. Repeat and take $3(5S^2)^{1/3}=\sigma$, whence $\log S=11.2587510$. Hence all positive integers $\leqq S$ are represented by $1_4, 2_2, 5_2)$. Finally, apply Lemma 5 with $T=0$, $t=\tau$, s replaced by v. Take $3(\tau v^2)^{1/3}=S$, whence $\log v\geqq 14.3976364$. Hence all positive integers $\leqq v$ are represented by $L=(1_4, 2_2, 5_2, \tau)$. By §§ 3, 4, all integers $\geqq K=\tau+Cp^{12}$ are represented by L. Since $\log K<13.7272$, $K<v$. The two results prove that all positive integers are represented by L.

Evidently we obtain a like conclusion if we start with B, D, E, F, or H instead of G, since ascent again begins with J.

For forms $(1_5, 5_2, t, l)$ obtained from C, we take $l=1$ and begin ascent

[6] *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 720.

from (1_6); but we can make only three ascents (instead of four as heretofore). The integers $N < 12{,}000$ which require 7, 8, or 9 cubes are listed by Jacobi.[7] It is readily verified that $N - 5$ is a sum of six cubes unless $N = 47{,}111{,}300$. But $111 - 5 \cdot 2^3 = 71$ and $300 - 5 \cdot 2^3 = 260$ are sums of five cubes; while $47 - 5 - 5$ is a sum of four. The table by Von Sterneck[8] shows that each integer > 8042 and $< 40{,}000$ is a sum of six cubes. Hence all integers $< 40{,}000$, except only 47, are represented by $(1_6, 5)$ while all are represented by $(1_6, 5_2)$. Taking $t = 5$ and making three ascents, we find that $(1_6, 5_3)$ represents all integers $< s$ if $\log s = 10.143227$. Our least C is 15.006868; for it, $\log (t + 11^{12}C) = 13.6730025 = L$.

But by use of Von Sterneck's table to 40,000, we can show that also all integers from 40,000 to 700,000 are sums of six cubes. The three ascents now lead to s, with $\log s > L$.

THEOREM 5. *The following* 6344 *forms are universal:* $(1_5, 2_2, t, l)$, $(1_3, 2_4, t, l)$, $(1_3, 2_2, 3_2, t, l)$, $(1_3, 2_2, 4_2, t, l)$, $(1_3, 2_2, 5_2, t, l)$, $(1_3, 2_2, 6_2, t, l)$, *where* t, l *have the values under* B, D, E, F, G, H, *respectively in* § 4. *Also,* $(1_6, 5_3)$.

6. There exist infinitely many primes $p \equiv 2 \pmod 3$ such that p is any of the four forms P in Lemma 3. If $0 < q \leqq 1$, take $p^6 \geqq g^3/q$. Then (22) holds if $C = l + w + q$. Evidently (23) fails if $l \geqq 6$, but holds if $l \leqq 5$ when p is sufficiently large. The same is true when $\nu = 3$ whence p^2 is replaced by p in (23). Hence the general theory in § 3 applies when $\nu = 3$ or 4, $1 \leqq l \leqq 5$, t arbitrary, and p sufficiently large.

Lemma 3 fails for the majority of forms f. For example, $x^2 + y^2 + 3z^2$ represents all positive integers $\neq 9^k(9r + 6)$; but $9 \equiv \gamma^3 + 6\mu \pmod{27}$ requires $\mu \equiv 6 \pmod 9$. Also, $x^2 + y^2 + 6z^2$ represents all $\neq 9^k(9r + 3)$, while $18 \equiv \gamma^3 + 6\mu \pmod{27}$ requires $\mu \equiv 3 \pmod 9$. Next, $x^2 + 3y^2 + 3z^2$ represents all $\neq 9^k(3r + 2)$, while $3 \equiv \gamma^3 + 6\mu \pmod 9$ requires $\mu \equiv 2 \pmod 3$. Finally, $x^2 + y^2 + 4z^2$ represents all except $8r + 3$ and $4^k(8r + 7)$, while $2 \equiv \gamma^3 + 6\mu \pmod 8$ requires $\mu \equiv 3 \pmod 4$.

Report will be made later on the result of using various primes p instead of our single p, on forms involving only eight cubes, and on the generalization of my papers on sums of values of cubic polynomials to sums of their products by arbitrary integers.

THE UNIVERSITY OF CHICAGO.

[7] Werke, VI, p. 323; *Journal für Mathematik*, vol. 42 (1851), p. 41.

[8] *Akademie Wissenschaften, Wien, Sitzungsberichte*, vol. 112, IIa (1903), pp. 1627-1666.

FUNDAMENTAL SYSTEMS OF UNITS IN NORMAL FIELDS.

By MARIE J. WEISS.

Recently Professor Latimer proved some interesting theorems[1] on the fundamental systems of units in algebraic fields cyclic with respect to the rational field. Here some of these theorems are extended to non-cyclic algebraic fields by a consideration of the integral group ring as a ring of operators for the group of units in the field.[2]

Let Ω be a normal extension of degree g of the rational field and G the Galois group of Ω. From the Galois theory it is known that the maximum real subfield Λ of Ω belongs to a subgroup H of order one or two according as Ω is real or imaginary. In either case denote the degree of Λ by $r+1$. In a normal field a unit η_1 can always be chosen in Λ such that η_1 and its conjugates generate a subgroup M of finite index m in the group of all units U of Ω.[3] Every r of the conjugates form a system of independent units in Ω.

The quotient group U/R, where R is the group of roots of unity in Ω, is invariant under G. Since M is generated by a complete set of conjugates under G, the quotient group M/R', R' being the subgroup of roots of unity found in M, is also invariant under G. When these two quotient groups have been written additively R and R' correspond to zero, and throughout the paper we shall so write them. Note that the classes of U/R containing units which form a fundamental system of units in Ω form an independent basis for U/R, while any r of those classes of M/R' which contain one of the conjugate units $\eta_1, \eta_2, \cdots, \eta_{r+1}$ form an independent basis for M/R'.

If ζ is any unit of Ω and S any element of G, by $S\zeta$ we shall mean that element obtained from ζ by performing the operation S. By $\sum_{i=1}^{g} a_i S_i \zeta$ we shall mean the unit $\prod_{i=1}^{g} (S_i\zeta)^{a_i}$. Hence the group ring with rational integral coefficients, which we shall denote by $[G]$, can be taken as a ring of operators[4] for the additive modules M/R' and U/R.

---

[1] C. G. Latimer, *American Journal of Mathematics*, vol. 56, no. 1 (1934), pp. 69-74.

[2] The author is indebted to the late Professor Emmy Noether for suggesting this method of treating the problem.

[3] H. Minkowski, *Goettinger Nachrichten* (1900), p. 90; H. Hasse, *Klassenkoerpertheorie, Anhang. Theorie der Einheiten.*

[4] E. Noether, *Mathematische Zeitschrift*, vol. 30, pp. 641-692.

We shall use a notation which will combine the cases Ω real and Ω imaginary. If in all that follows the sum $1/2\Sigma H$, the sum being over the elements of H, is replaced by the identity E, the imaginary case reduces to the real case. Let (G) denote the group ring with rational coefficients. The left ideal $(G)\Sigma H$ can be written as the direct sum of the two left ideals $(G)E_1$ and $(G)E_2$, where $E_1 = 1/g\Sigma G$ and $E_2 = 1/2\Sigma H - 1/g\Sigma G$ are idempotents. Write the group $G = H + S_2H + \cdots + S_{r+1}H$. Using the integral group ring as a ring of operators, we shall set up an operator isomorphism between the $[G]$-left module $[G]mE_2 = [G]m/2\Sigma HE_2$, which we shall denote by $\mathfrak{M}$, and the additive module M/R'. Note that $1/2\Sigma H\eta_1 = \eta_1$, and label the conjugates of η_1 so that $S_i\, 1/2\Sigma H\eta_1 = \eta_i$, $i = 2, 3, \cdots, r+1$. Further denote the class of M/R' containing ζ by ζ^*. If we let

$$m/2\Sigma HE_2 = mE_2 \rightarrow \eta_1^*$$

and

$$m/2\, S_i\Sigma HE_2 = mS_iE_2 \rightarrow \eta_i^* \qquad (i = 2, 3, \cdots, r+1),$$

we have the desired operator isomorphism. Now

$$m/2 \sum_{i=1}^{r+1} S_i\Sigma HE_2 = 0 \rightarrow \sum_{i=1}^{r+1} \eta_i^* = 0$$

and

$$m/2 \sum_{i=1}^{r+1} a_iS_i\Sigma HE_2 \rightarrow \sum_{i=1}^{r+1} a_i\eta_i^*$$

equals zero if and only if the a_i are equal. Note that if

$$w_1 = m/2 \sum_{i=1}^{r+1} a_iS_i\Sigma HE_2 \rightarrow \zeta^*$$

and if

$$w_2 = m/2 \sum_{i=1}^{r+1} b_iS_i\Sigma HE_2 \rightarrow \zeta^*$$

then

$$w_1 - w_2 = m/2 \sum_{i=1}^{r+1} (a_i - b_i)S_i\Sigma HE_2 = 0,$$

and $a_i - b_i = k$, thus showing that w_1 and w_2 differ only by zero. We conclude that only r of the elements $m/2\Sigma HE_2$, $m/2S_2\Sigma HE_2, \cdots, m/2S_{r+1}\Sigma HE_2$ are linearly independent, and any r of them may be taken as an independent basis of $\mathfrak{M}$.

Since M is of finite index m in U, the m-th power of every unit in Ω is found in M. From the operator isomorphism we have established, we see that the image of every class of U/R is found in the $[G]$-left module $[G]E_2$.

If $\zeta_1, \zeta_2, \cdots, \zeta_r$ form a fundamental system of units in U, we have $u'_i \rightarrow m\xi_i^*$, u'_i an element of $\mathfrak{M}$, $i = 1, 2, \cdots, r$. Hence $u'_i/m = u_i \rightarrow \xi_i^*$. Thus all linear forms of the u_i with integral coefficients generate a module $\mathfrak{U}$ which is operator isomorphic to the additive module U/R, the integral group ring being the ring of operators. Note that if the elements $v_1, v_2, \cdots, v_r$ are elements of a second basis of $\mathfrak{U}$, then $(v_1, v_2, \cdots, v_r) = (u_1, u_2, \cdots, u_r)A$, where A is a matrix whose determinant equals ± 1, but then

$$(\zeta_1', \zeta_2', \cdots, \zeta_r') = (\zeta_1, \zeta_2, \cdots, \zeta_r)A,$$

and $\zeta_1', \zeta_2', \cdots, \zeta_r'$ form a fundamental system of units of U. Thus we have a one to one correspondence between the independent bases of $\mathfrak{U}$ and the fundamental systems of units in U.

The above operator isomorphism enables us to establish a criterion for the existence of a fundamental system of units composed of conjugates. If the modules $\mathfrak{M}$ and $\mathfrak{U}$ are operator isomorphic under the integral group ring, we see that a fundamental system of units composed of conjugates may be chosen. On the other hand if a fundamental system of units composed of conjugates exists, $\mathfrak{M}$ and $\mathfrak{U}$ are operator isomorphic. Hence we may state

THEOREM 1. *A necessary and sufficient condition for the existence of a fundamental system of units composed of conjugates is that the modules* $\mathfrak{M}$ *and* $\mathfrak{U}$ *be operator isomorphic under the integral group ring.*

Note that if Ω is real or if the subgroup H is normal, $\mathfrak{M}$ is a two sided $[G]$-module, generated by a single element. Hence in these cases our criterion for the existence of a fundamental system of units composed of conjugates is analogous to the one given by Latimer.[5]

We may also draw some conclusions concerning the existence of a fundamental system of real units. If Ω is imaginary and a fundamental system of real units exists, the conjugate imaginary $\bar{\eta}$ of a unit η equals $\rho^a\eta$, where ρ is a root of unity generating R. In this case every class of U/R is invariant under H. However, from the operator isomorphism we have set up, we see that every class of U/R is left invariant by H, if and only if H is invariant. Moreover, if every class of U/R is invariant under H, a unit and its conjugate imaginary are found in the same class. Hence we may seek fundamental systems of real units only in those fields whose maximum real subfield is normal.

THEOREM 2. *A necessary condition for the existence of a fundamental*

[5] C. G. Latimer, *loc. cit.*, Theorem 2.

system of real units in an imaginary field Ω normal with respect to the rational field is the invariance of the subgroup H of the Galois group G to which the maximum real subfield of Ω belongs. If H is invariant the conjugate imaginary of a unit η of Ω equals $\rho^a\eta$, where ρ is a root of unity.

Note that the conditions of the above theorem are satisfied in Abelian fields. *In all that follows we shall assume that H is an invariant subgroup of G.*

In case Ω is an imaginary cyclic field, the above theorem enables us to prove without using a matric representation of the group of units Latimer's theorem that a fundamental system of real units exists.[6] Let $\zeta_1, \zeta_2, \cdots, \zeta_r$ be a fundamental system of units. Since G is cyclic, H is invariant and $\bar{\zeta}_i = \rho^{a_i}\zeta_i$. Let $H = \{E, S^{r+1}\}$, S being a generating element of G. If $a_i = 2b_i$, the b_i integers, we choose the units $\rho^{b_i}\zeta_i$ as the elements of a fundamental system. These are real since $S^{r+1}(\rho^{b_i}\zeta_i) = \rho^{-b_i}\bar{\zeta}_i = \rho^{-b_i}\rho^{2b_i}\zeta_i = \rho^{b_i}\zeta_i$. We shall show that each a_i is even. Recall that R is generated by a primitive even root of unity. Let $S\rho = \rho^h$, then $S^{r+1}\rho = \rho^{h^{r+1}} = \rho^{-1}$ from which we conclude that h is odd. Now letting $S\zeta_i = \zeta_i^{(1)}$ and $S^v\zeta_i = \zeta_i^{(v)}$, we have

$$\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)}\zeta_i^{(r+1)} \cdots \zeta_i^{(2r+1)} = 1,$$

where $\zeta_i^{(r+j)}$ is the conjugate imaginary of $\zeta_i^{(j-1)}$ and $\zeta_i^{(0)} = \zeta_i$. Using the relation $\bar{\zeta}_i = \rho^{a_i}\zeta_i$, we obtain

$$(\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)})^2 = \rho^l$$

and since $\zeta_i \cdots \zeta_i^{(r)}$ is in Ω as well as its square

$$\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)} = \rho^k, \qquad k \text{ an integer.}$$

Operating by S on both sides of this last equation, we obtain

$$\zeta_i^{(1)}\zeta_i^{(2)} \cdots \zeta_i^{(r+1)} = \rho^{hk}$$

which gives

$$\zeta_i^{(r+1)} = \rho^{k(h-1)}\zeta_i.$$

Since h is odd we have proved our proposition.

We seek the image of those classes of U/R which contain units from a subfield Γ of Ω. Let the elements of K, the subgroup of index k to which the subfield Γ belongs, be $K_1 = E, K_2, \cdots, K_k$. Write G in terms of its v cosets with respect to the subgroup generated by K and H. If K con-

[6] C. G. Latimer, *loc. cit.*, Theorem 1.

tains H, $v = k$, while if K does not contain H, $v = k/2$. Let η^* be a class of U/R which contains a unit from the subfield Γ, and its image in $[G]E_2$ be

$$w = 1/2(\Sigma a_i K_i \Sigma H + \Sigma b_i K_i \Sigma H S_2 + \cdots + \Sigma v_i K_i \Sigma H S_v) E_2,$$

$i = 1, 2, \cdots, s$, where $s = h/2$ or h according as K does or does not contain H. Since $K\eta^* = \eta^*$, $Kw = w$, K any element of the group K. Equating coefficients in Kw and w, we have $a_i - a_k = t$, $b_i - b_k = t, \cdots, v_i - v_k = t$, the subscripts k being a permutation of the subscripts i. Hence

$$\sum_{i=1}^{s} (a_i - a_k) = st = 0, \quad t = 0, \text{ and } a_i = a_k.$$

By operating in turn with each element of K, we see that the $a_i, b_i, \cdots, v_i$, respectively, are sets of equal integers. Thus to each class of U/R which contains a unit from the subfield Γ corresponds an element of the form:

$$1/2(a_1 \Sigma K_i \Sigma H + a_2 \Sigma K_i \Sigma H S_2 + \cdots + a_v \Sigma K_i \Sigma H S_v) E_2, \quad (i = 1, \cdots, s).$$

These elements are contained in a module $\mathfrak{R}$ which is a right-sided $[G]$-module. If we regard the v sums $1/2 \sum_{i=1}^{s} K_i \Sigma H S_j E_2$ as elements, we see that every $v - 1$ of them form an independent basis for the module $\mathfrak{R}$. Note that $\mathfrak{R}$ may be generated by a single element $1/2 \Sigma K_i \Sigma H E_2$, by multiplication on the right by $[G]$. Further if the subgroup K is invariant, $\mathfrak{R}$ is a two-sided $[G]$-module. We may summarize as follows:

THEOREM 3. *The images of those classes of U/R which contain units from the subfield belonging to the subgroup K of G are found in the right-sided $[G]$-module* $(1/2 \sum_{i=1}^{s} K_i \Sigma H E_2)[G]$.

We now apply the theory developed thus far to the case where G is the direct product of two subgroups S and T of relatively prime orders s and t, respectively. Hence one subgroup contains H, say S does. We need to bear in mind that we are assuming H an invariant subgroup of G. For convenience let $s/2 = q$. Now $(G)E_2$ contains the following idempotents:

$I_1 = 1/s \Sigma S E_2 = 1/s \Sigma S - 1/g \Sigma G,$
$I_2 = 1/t \Sigma T E_2 = 1/(2t) \Sigma T \Sigma H - 1/g \Sigma G,$ and
$I_3 = (1/2 \Sigma H - 1/s \Sigma S - 1/t \Sigma T) E_2 = 1/2 \Sigma H - 1/s \Sigma S - 1/(2t) \Sigma T \Sigma H + 1/g \Sigma G,$

such that $I_1 + I_2 + I_3 = E_2$, $I_i I_j = 0$, $i \neq j$, $I_i^2 = I_i$, $I_i E_2 = E_2 I_i = I_i$, $i = 1, 2, 3$. Hence $(G)E_2$ is the direct sum of $(G)I_1$, $(G)I_2$, and $(G)I_3$. Now consider the submodule $\mathfrak{W}$ formed by the direct sum of $\mathfrak{R} = [G]I_1$, $\mathfrak{S} = [G]I_2$, and $\mathfrak{T} = [G]I_3$. $\mathfrak{R}$ has an independent basis of $r_1 = t - 1$ of the t elements $I_1, T_2 I_1, \cdots, T_t I_1$, $\mathfrak{S}$, an independent basis of $r_2 = q - 1$ of

the q elements $1/2\Sigma HI_2$, $1/2\Sigma HS_2I_2, \cdots, 1/2\Sigma HS_qI_2$, and $\mathfrak{T}$ an independent basis of $r_3 = r - t - q + 2$ elements. Since $\mathfrak{U}$ is a submodule of $\mathfrak{W}$, its basis can be expressed in the form:

$$\begin{aligned} u_i &= x_i && (i = 1, 2, \cdots, r_1), \\ u_i &= x_i' + y_i && (i = r_1 + 1, \cdots, r_1 + r_2), \\ u_i &= x_i'' + y_i' + z_i && (i = r_1 + r_2 + 1, \cdots, r), \end{aligned}$$

where the x_i, x_i', x_i'' are linear combinations with integral coefficients of the base elements of $\mathfrak{R}$ and y_i and y_i' similar functions of the base elements of $\mathfrak{S}$, while the z_i are such functions of the base elements of $\mathfrak{T}$. Or, in other words, if we consider the matrix of the base elements of $\mathfrak{W}$, we may take zeros above the principal diagonal. Since the module $\mathfrak{U}$ admits $[G]$ as a ring of operators, we have,

$$\Sigma S u_i = s x_i'$$

and

$$t(E - 1/t\Sigma T)u_i = t x_i' \qquad (i = r_1 + 1, \cdots, r_1 + r_2).$$

Since the integers s and t are relatively prime, each of the x_i' is an element of $\mathfrak{U}$. Thus we may take a basis of $\mathfrak{U}$ in the above form with the elements x_i' replaced by zero.

As we have previously seen to an independent basis u_i of $\mathfrak{U}$ corresponds a fundamental system of units $\zeta_1, \zeta_2, \cdots, \zeta_r$ of Ω. Since the first r_1 of the u_i are invariant under the elements of the group S, while the next r_2 are invariant under the elements of T, the first r_1 of the ζ_i lie in the subfield Γ_1 which belongs to the subgroup S, and the next r_2 of the ζ_i lie in the subfield Γ_2 which belongs to the subgroup T. Since the ζ_i form a fundamental system of units in Ω, the first r_1 of the ζ_i form such a fundamental system in Γ_1 and the next r_2 form such a system in Γ_2. Hence we have

THEOREM 4. *Let Ω be an absolutely normal algebraic field in which the maximum real subfield is normal in case Ω is imaginary. If Ω is obtained by the composition of two subfields Γ_1 and Γ_2, of relatively prime degrees, Ω contains a fundamental system of units such that certain of them form a fundamental system in Γ_1 and certain others form such a system in Γ_2.*

If Γ_2 is again obtained by the composition of two fields of relatively prime degrees, we can apply the previous theorem to Γ_2 instead of Ω, using the quotient group G/T instead of G, and thus obtain a fundamental system of units of Γ_2 such that it contains fundamental systems from each of the two subfields which generate Γ_2. An Abelian group can always be written as the direct product of subgroups whose orders are powers of distinct primes. Hence the theorem may always be applied to Abelian fields.

BRYN MAWR COLLEGE.

A CALCULUS OF SEQUENCES.

By MORGAN WARD.

I. *Introduction.*

1. I propose in this paper to give a generalization of a large portion of the formal parts of algebraic analysis and the calculus of finite differences. The generalization consists in systematically replacing the ordinary binomial coefficient $n(n-1)\cdots(n-r+1)/1\cdot 2\cdots r$ by a "binomial coefficient to the base (u)," $[n,r]=u_n\cdot u_{n-1}\cdots u_{n-r+1}/u_1\cdot u_2\cdots u_r$ where (u): $u_0, u_1, u_2, \cdots$ is a fixed sequence of complex numbers subject to the restrictions $u_0=0$; $u_1=1$; $u_n\neq 0$, $n>1$. The exponential function for example is replaced by the formal series $1+\sum_{n=1}^{\infty} x^n/u_1\cdot u_2\cdots u_n$, differentiation by an operation which throws x^n into $[n,1]x^{n-1}$, and differencing by an operation which throws x^n into $\sum_{r=1}^{n}[n,r]x^{n-r}$.

The formula $n^r=(1+1+\cdots+1)^r=\sum_{(s)} r!/s_1!s_2!\cdots s_n!$ guides us in replacing the powers of rational integers where necessary by sums of multinomial coefficients to the base (u); for example, $\sum_{s=0}^{r}[n,r]$ may replace 2^r. We are thus enabled to generalize successfully a great variety of formulas involving the exponential functions, the Bernoulli numbers and polynomials.

2. In a series of papers which have appeared during the past thirty years [1],[1] F. H. Jackson has developed a somewhat similar extension of elementary analysis for the particular sequence (u) in which $u_n=(q^n-1)/(q-1)$ q a fixed complex number, $|q|\neq 1$. His results are based essentially on an identity of Euler's [2]:

$$(2.1)\qquad (x+y)(x+qy)\cdots(x+q^{n-1}y)=\sum_{r=0}^{n}[n,r]q^{r(r-1)/2}x^{n-r}y^r.$$

In effect he replaces the ordinary binomial coefficient by $[n,r]q^{r(r-1)/2}$. But the presence of this power of q introduces a lack of symmetry in his formulas,

[1] The numbers in square brackets refer to references at the end of the paper. Jackson wrote in all over thirty papers on this subject. We have listed only those directly connected with the present paper.

and leads to certain complications in defining the exponential functions.[2] He was not led furthermore to consider the developments of the calculus of finite differences which we give here. These appear to be among the most striking results of the entire theory.

The present work originated in an (unsuccessful) attempt to frame a definition of the Bernoulli numbers in Jackson's calculus to which the Staudt-von Klausen theorem might apply.

II. *Formal theory.*

3. Let

$$(u): \quad u_0 = 0, \quad u_1 = 1, \quad u_2, \cdots, u_n, \cdots$$

be a fixed sequence of complex numbers subject for the present to the single restriction $u_n \neq 0$, $n > 1$. For convenience, we shall write $[n]$ for u_n. We define:

$[n]!$ to be 1 if $n = 0$, and $[n]\,[n-1] \cdots [1]$ if $n > 0$;
$[n, r]$ to be $[n]!/[r]![n-r]!$

where n, r are positive integers,[3] and $n \geqq r$. Then

$$[n, 0] = 1, \quad [n, 1] = [n], \quad [n, n-r] = [n, r].$$

We shall call $[n, r]$ a binomial coefficient to the base (u), or simply a *basic*[4] binomial coefficient.

We write $(x+y)^n$ for the polynomial $\sum_{r=0}^{n} [n, r] x^{n-r} y^r$. It is evident that

$$(x+y)^0 = 1, \quad (x+y)^1 = x+y, \quad (x+0)^n = x^n$$
$$(cx+cy)^n = c^n (x+y)^n, \quad (x+y)^n = (y+x)^n.$$

From the identities[5]

$$(x-y)^{2n+1} = \sum_{r=0}^{n} (-1)^r [2n+1, r] x^r y^r (x^{2n+1-2r} - y^{2n+1-r})$$

$$(x-y)^{2n} = \sum_{r=0}^{n-1} (-1)^r [2n, r] x^r y^r (x^{2n-2r} + y^{2n-2r}) + (-1)^n [2n, n] x^n y^n$$

---

[2] It is necessary to consider not only the series $1 + \sum_{1}^{\infty} x^n / u_1 u_2 \cdots u_n$ as an analogue of the exponential, but also the series $1 + \sum_{1}^{\infty} q^{n(n-1)/2} x^n / u_1 u_2 \cdots u_n$ with a corresponding complexity in the theory of the trigonometric functions.

[3] We count zero as a positive integer.

[4] This convenient terminology is due to F. H. Jackson.

[5] We write $(x-y)^n$ for $(x + (-y))^n = \sum_{r=0}^{n} [n, r] x^{n-r} (-y)^r$.

we see that

$$(x - x)^{2n+1} = 0; \quad (n = 0, 1, 2, \cdots).$$

On the other hand,

$$(x - x)^{2n} = x^{2n}(1 - 1)^{2n} = x^{2n}\{2\sum_{r=0}^{n-1}(-1)^r[2n, r] + (-1)^n[2n, n]\}$$

generally does *not* vanish. A sequence (u) such that

$$(1 - 1)^{2n} = 0, \ (n = 1, 2, 3, \cdots)$$

will be said to be *normal.*

4. More generally, we define

$$(x_1 + x_2 + \cdots + x_k)^n \text{ to be } \sum_{(s)} \frac{[n]!}{[s_1]! \cdots [s_k]!} x_1^{s_1} \cdots x_k^{s_k}$$

where the summation is over all integers s satisfying the conditions

$$s_1 + s_2 + \cdots + s_k = n, \qquad 0 \leqq s_i \leqq n.$$

If we denote this polynomial by $P_{kn}(x) = P_{kn}(x_1, x_2, \cdots, x_k)$ then it is a symmetric function of its k arguments, and if c is any constant, then

$$P_{kn}(cx_1, cx_2, \cdots, cx_k) = c^n P_{kn}(x_1, x_2, \cdots, x_k).$$

Furthermore,

$$P_{k+1n}(x) = P_{kn}(x_1, x_2, \cdots, x_{k-1}, x_k + x_{k+1}). \tag{4.1}$$

For consider $P_{2n}(x) = (x_1 + x_2)^n$. We have

$$\begin{aligned}
(x_1 + x_2)^n &= \sum_{t=0}^{n} \frac{[n]!}{[n-t]!\,[t]!} x_1^{n-t} x_2^{t} \\
(x_1 + (x_2 + x_3))^n &= \sum_{t=0}^{n} \frac{[n]!}{[n-t]!\,[t]!} x_1^{n-t}(x_2 + x_3)^t \\
&= \sum_{t=0}^{n} \sum_{r=0}^{t} \frac{[n]!\,[t]!}{[n-t]!\,[t]!\,[t-r]!\,[r]!} x_1^{n-t} x_2^{t-r} x_3^{r} \\
&= \sum_{(s)} \frac{[n]!}{[s_1]!\,[s_2]!\,[s_3]!} x_1^{s_1} x_2^{s_2} x_3^{s_3}, \\
&\qquad\qquad s_1 + s_2 + s_3 = n, \ 0 \leqq s_i \leqq n \\
&= (x_1 + x_2 + x_3)^n = P_{3n}(x).
\end{aligned}$$

Hence (4.1) is true for $k = 2$. Its validity for any value of k follows by an easy induction.

It is evident that formula (4.1) can be extended so as to express $P_{k+ln}(x)$ in terms of $P_{kn}(x)$ in various ways. For example,

$$P_{4n}(x_1, x_2, x_3, x_4) = P_{2n}(x_1 + x_2, x_3 + x_4).$$

5. The numerical values of the polynomials $P_{kn}(x)$ when all of the arguments x_i are equal to plus one play an important rôle in the developments which are to follow. We shall write $\overline{k}^n$ for the number

$$P_{kn}(1, 1, \cdots, 1) = (1 + 1 + \cdots + 1)^n = \sum_{(s)} \frac{[n]!}{[s_1]! \cdots [s_k]!}.$$

We see from the formulas of section 4 that

$$\overline{3}^n = (\overline{2} + \overline{1})^n, \quad \overline{4}^n = (\overline{2} + \overline{2})^n = (\overline{3} + \overline{1})^n.$$

It is easily shown by induction that we have quite generally

$$\overline{r + s}^n = (\overline{r} + \overline{s})^n \tag{5.1}$$

where r and s are any positive integers.

If furthermore the sequence (u) is normal (section 3) then we can show by induction that (5.1) holds for any integral values of r and s. A somewhat longer induction establishes the formula

$$\overline{m_1 + m_2 + \cdots + m_t}^n = (\overline{m}_1 + \overline{m}_2 + \cdots + \overline{m}_t)^n \tag{5.2}$$

where if (u) is normal, $m_1, \cdots, m_t$ are any integers, but if (u) is not normal, the integers are to be positive.

In case (u) is normal, there is no gain in generality in replacing some of the plus signs in formula (5.2) by minus signs because we can show that

$$(\overline{r} - \overline{s})^n = (\overline{r} + \overline{-s})^n. \tag{5.3}$$

6. If $F(x)$ denotes the formal power series

$$F(x) = \sum_{n=0}^{\infty} c_n x^n, \tag{6.1}$$

we define $F(x + y)$ to mean the series

$$\sum_{n=0}^{\infty} c_n (x + y)^n = \sum_{n=0}^{\infty} \sum_{r=0}^{n} c_n [n, r] x^{n-r} y^r.$$

In like manner

$$F(x_1 + x_2 + \cdots + x_k) = \sum_{n=0}^{\infty} c_n (x_1 + x_2 + \cdots + x_k)^n = \sum_{n=0}^{\infty} c_n P_{kn}(x). \tag{6.2}$$

We have furthermore formal identities of the type

$$F(x_1 + x_2) = F(x_2 + x_1),$$
$$F(x_1 + x_2 + x_3) = F(x_1 + (x_2 + x_3))$$
$$F(x_1 + x_2 + x_3 + x_4) = F((x_1 + x_2) + (x_3 + x_4))$$

since the like identities hold for the polynomials $P_{kn}(x)$.

If in the series (6.2) we make all the arguments x_i equal to x, the right side becomes $\sum_{n=0}^{\infty} c_n(x + x + \cdots + x)^n = \sum_{n=0}^{\infty} c_n \bar{k}^n x^n$. We shall accordingly denote the resulting series by $F(\bar{k}x)$. It is obvious then from the formula (5.2) that

$$(6.3) \quad F(\overline{m_1 + m_2 + \cdots + m_t}x) = F(\bar{m}_1 x + \bar{m}_2 x + \cdots + \bar{m}_t x)$$

for suitably restricted integers m_i.

Let $F(x)$, $G(x)$, $H(x)$ be three formal power series in x. Then the following theorem is easily seen to be true.

THEOREM 6.1. *If $F(x) = G(x) \pm H(x)$ and m, n are any positive integers, then $F(\bar{m}x) = G(\bar{m}x) \pm H(\bar{m}x)$ and*

$$F(\bar{m}x + \bar{n}y) = G(\bar{m}x + \bar{n}y) \pm H(\bar{m}x + \bar{n}y).$$

7. We next define an operator $D = D_x$ which transforms the formal power series (6.1) into

$$(7.1) \qquad F'(x) = DF(x) = \sum_{n=0}^{\infty} [n] c_n x^{n-1}.$$

In particular then, $Dx^n = [n]x^{n-1}$. The operator D is easily shown to be linear and distributive, and it converts a polynomial of degree n in x into one of degree $n - 1$.

If we define $F^{(r)}(x) = D^r F(x)$ recursively by $F^{(r+1)}(x) = DF^{(r)}(x)$; $F^{(0)}(x) = F(x)$, it easily follows that

$$\frac{F^{(r)}(x)}{[r]!} = \sum_{n=r}^{\infty} [n, r] c_n x^{n-r}.$$

The expansion $F(x + y) = \sum_{n=0}^{\infty} c_n(x + y)^n$ is formally replaceable by

$$(7.2) \qquad F(x + y) = \sum_{n=0}^{\infty} \frac{F^{(n)}(x)}{[n]!} y^n.$$

We shall refer to (7.2) as *Taylor's formula* for the base (u).

Finally, we note that

$$D_x{}^r F(x + y) = F^{(r)}(x + y).$$

8. As a simple concrete example of such an operator D, let us assume that the sequence (u) is a linear recurring series of order k whose associated polynomial $x^k - a_1x^{k-1} - \cdots - a_k$ has k distinct roots $\alpha_1, \alpha_2, \cdots, \alpha_k$. Then u_n is of the form

$$u_n = \beta_1\alpha_1{}^n + \beta_2\alpha_2{}^n + \cdots + \beta_k\alpha_k{}^n$$

where the constants α and β are subject to the conditions

$$\beta_1 + \beta_2 + \cdots + \beta_k = 0, \quad \beta_1\alpha_1 + \beta_2\alpha_2 + \cdots + \beta_k\alpha_k = 1, \quad u_n \neq 0, n > 1.$$

It is obvious then that

$$DF(x) = (\beta_1F(\alpha_1x) + \cdots + \beta_kF(\alpha_kx))/(\beta_1\alpha_1x + \cdots + \beta_k\alpha_kx).$$

This operator can therefore be applied to any function of x regular at $x = 0$, and transforms it into another function regular at $x = 0$.

In particular, if $k = 2$, $\alpha_1 = q$, $\alpha_2 = 1$, $\beta_1 = (q-1)^{-1}$, $\beta_2 = -\beta_1$, where q is not a root of unity,

$$DF(x) = \frac{F(qx) - F(x)}{qx - x}$$

is the operation of q-differencing.

In case some of the roots α of the polynomial associated with the recurrence relation are repeated, a similar but more complicated formula for $DF(x)$ may be given which involves both $F(x)$ and its ordinary derivatives. For example, if $k = 2$ and $\alpha_1 = \alpha_2 \neq 0$, $u_n = n\alpha_1{}^{n-1}$ and $DF(x) = \frac{1}{\alpha_1}\frac{dF(\alpha_1x)}{dx}$.

III. *The exponential and trigonometric functions.*

9. We shall now assume that the sequence (u) is chosen in such a manner that the series

$$\mathcal{E}(x) = \sum_{n=0}^{\infty} \frac{x^n}{[n]!} \tag{9.1}$$

is convergent in the neighborhood of $x = 0$. It accordingly is an element of an analytic function of x which we shall call the *basic exponential*. There exists then a positive number ρ such that the series (9.1) converges absolutely within the circle $|x| = \rho$.

The basic exponential has the following properties for sufficiently small absolute values of its arguments x, y, x_i:

$$D\mathcal{E}(x) = \mathcal{E}(x), \quad \mathcal{E}^{(n)}(cx) = c^n\mathcal{E}(cx), c \text{ a constant}, \tag{9.2}$$

$$\mathcal{E}(x+y) = \mathcal{E}(x)\mathcal{E}(y), \tag{9.21}$$

$$\mathcal{E}(x_1 + x_2 + \cdots + x_k) = \mathcal{E}(x_1)\mathcal{E}(x_2)\cdots\mathcal{E}(x_k). \tag{9.22}$$

Consider for example the formula (9.21). That it is formally true is immediately obvious from the basic Taylor's formula (7.2). For

$$\mathcal{E}(x+y) = \sum_{n=0}^{\infty} \frac{\mathcal{E}^{(n)}(x) y^n}{[n]!} = \mathcal{E}(x)\mathcal{E}(y)$$

since by (9.2), $\mathcal{E}^{(n)}(x) = \mathcal{E}(x)$.

But the series

$$\sum_{n=0}^{\infty} \frac{(x+y)^n}{[n]!} = \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{x^{n-r} y^r}{[n-r]!\,[r]!}$$

is in fact the Cauchy product of the series $\Sigma \frac{x^n}{[n]!}, \Sigma \frac{y^n}{[n]!}$ so that the formula is actually true provided the latter two series are both absolutely convergent. And by our initial hypothesis, both series converge absolutely if $|x| < \rho$, $|y| < \rho$.

10. The trigonometric and hyperbolic functions are defined by Euler's formulas:

$$\begin{aligned} &\sin(x) = \frac{\mathcal{E}(ix) - \mathcal{E}(-ix)}{2i}, \quad \cos(x) = \frac{\mathcal{E}(ix) + \mathcal{E}(-ix)}{2}, \\ &\sinh(x) = -i \sin(ix), \quad \cosh(x) = \cos(ix). \end{aligned} \tag{10.1}$$

Among the many formal analogies with the ordinary trigonometric functions, we shall merely note here:

$$\begin{aligned} &\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y), \\ &\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y), \\ &D\sin(x) = \cos(x), \qquad D\cos(x) = -\sin(x). \end{aligned}$$

As a consequence of the last two formulas, we see that both $\sin(x)$ and $\cos(x)$ satisfy the basic differential equation $y^{(2)}(x) + y(x) = 0$.

On the other hand

$$\sin^2(x) + \cos^2(x) = \mathcal{E}(ix)\mathcal{E}(-ix) \tag{10.2}$$

and in general, $\mathcal{E}(ix)\mathcal{E}(-ix) \neq 1$.

The remaining trigonometric and hyperbolic functions are defined in terms of the basic sine and cosine as in the ordinary case.

11. It is possible to give analogues of De Moivre's and Simpson's formulas. For in formula (9.22), take $k = n$ and let $x_1 = x_2 = \cdots = x_n = i\theta$. Then with the notation explained in section 6,

$$\mathcal{E}(\bar{n} i\theta) = (\mathcal{E}(i\theta))^n. \tag{11.1}$$

Hence we obtain from the formulas (10.1) and theorem 6.1 the basic form of De Moivre's formula,[6] $\cos(\bar{n}\theta) + i\sin(\bar{n}\theta) = (\cos(\theta) + i\sin(\theta))^n$.

THEOREM 11.1. *The sequence* (u) *is normal when and only when* $\mathcal{E}(x)\mathcal{E}(-x) = 1$ *or when and only when* $\sin^2(x) + \cos^2(x) = 1$.

For by formula (9.21), and the previous definitions, if $|x| < \rho$,

$$\mathcal{E}(x)\mathcal{E}(-x) = \mathcal{E}(x-x) = \sum_{n=0}^{\infty} \frac{(x-x)^n}{[n]!} = \sum_{n=0}^{\infty} \frac{(1-1)^n}{[n]!} x^n.$$

Now we have seen in section 3 that $(1-1)^{2n+1} = 0$, $(n = 0, 1, 2, \cdots)$. Therefore

$$\mathcal{E}(x)\mathcal{E}(-x) = 1 + \sum_{n=0}^{\infty} \frac{(1-1)^{2n}}{[2n]!} x^{2n}.$$

Hence the first part of the theorem follows. The second part of the theorem is an immediate consequence of formula (10.2).

Let us assume now that (u) is normal. We see from formula (11.1) that

$$\begin{aligned} \mathcal{E}(\overline{n+2}i\theta) &= \mathcal{E}(i\theta)\mathcal{E}(\overline{n+1}i\theta), \\ \mathcal{E}(i\theta)\mathcal{E}(\bar{n}i\theta) &= \mathcal{E}(\overline{n+1}i\theta). \end{aligned} \tag{11.3}$$

But by theorem 11.1, $\mathcal{E}(-i\theta)\mathcal{E}(i\theta) = 1$. Therefore this last equation may be written

$$\mathcal{E}(\bar{n}i\theta) = \mathcal{E}(-i\theta)\mathcal{E}(\overline{n+1}i\theta). \tag{11.31}$$

On adding and subtracting the two formulas (11.3), (11.31) and applying (10.1) and theorem (6.1), we obtain the basic Simpson's formulas:

$$\begin{aligned} \cos(\overline{n+2}\theta) &= 2\cos(\theta)\cos(\overline{n+1}\theta) - \cos(\bar{n}\theta), \\ \sin(\overline{n+2}\theta) &= 2\cos(\theta)\sin(\overline{n+1}\theta) - \sin(\bar{n}\theta). \end{aligned}$$

12. In order that the results of the previous section may have more than a purely formal significance, it is necessary to show that we can choose the sequence (u) so that (u) is normal and so that $E(\bar{n}x)$ is an entire function of x for any integer n. Since $E(\overline{-n}x) = E(-\bar{n}x)$, $E(\bar{n}x) = (E(\bar{1}x))^n$, $E(\bar{1}x) = E(x)$, we need only consider the case when $n = +1$.

Now it is easy to show that the most general solution of the functional equation

$$\Phi(x)\Phi(-x) = 1 \tag{12.1}$$

[6] It should be noted here that $(\cos(\theta) + i\sin(\theta))^n$ stands for the product $(\cos(\theta) + i\sin(\theta))(\cos(\theta) + i\sin(\theta))\cdots$ taken to n factors, and not for the result of substituting $\cos(\theta)$ for x_1, and $i\sin(\theta)$ for x_2 in the polynomial $P_{2n}(x) = (x_1 + x_2)^n$.

which is regular at the origin is of the form

$$\Phi(x) = \pm \exp\,(x\Psi(x^2)) \tag{12.2}$$

where $\Psi(x)$ is regular at the origin. But by theorem 11.1, (u) is normal when and only when $\mathcal{E}(x)$ is a solution of (12.1). Since $\mathcal{E}(x)$ was assumed to be regular at the origin, $\mathcal{E}(x)$ must be of the form (12.2) where $\Psi(x)$ is an entire function of x. We must also satisfy the conditions[7]

$$\mathcal{E}(0) = 1, \quad \mathcal{E}'(0) = 1, \quad \frac{\mathcal{E}^{(n)}(0)}{n!} = \frac{1}{u_1 u_2 \cdots u_n} \neq 0,$$

as then $u_n = n\mathcal{E}^{(n-1)}(0)/\mathcal{E}^{(n)}(0) \neq 0$ $(n = 1, 2, \cdots)$ and $u_1 = 1$.

It will therefore suffice to choose for $\Psi(x)$ an entire function $G(x)$ with a series expansion of the form $G(x) = 1 + \sum_{n=1}^{\infty} g_n x^n$ where the quantities g_n are all real and non-negative. The ordinary case ensues on taking all the quantities g_n equal to zero.

13. If we assume that $\mathcal{E}(x)$ is an entire function satisfying the condition $\mathcal{E}(x)\mathcal{E}(-x) = 1$, we can generalize the periodic properties of the exponential function. For since $\mathcal{E}(x)$ never vanishes, by Picard's theorem there exists a complex number $\lambda \neq 0$ such that $\mathcal{E}(\lambda) = 1$. But then if n is a positive integer,

$$\mathcal{E}(x + \bar{n}\lambda) = \mathcal{E}(x)\mathcal{E}(\bar{n}\lambda) = \mathcal{E}(x)\,(\mathcal{E}(\lambda))^n = \mathcal{E}(x),$$
$$\mathcal{E}(x) = \mathcal{E}(x - \bar{n}\lambda + \bar{n}\lambda) = \mathcal{E}(x - \bar{n}\lambda)\mathcal{E}(\bar{n}\lambda) = \mathcal{E}(x - \bar{n}\lambda).$$

We have therefore proved the following theorem.

THEOREM 13.1. *If (u) is a normal sequence so chosen that the basic exponential function $\mathcal{E}(x)$ is an entire function of x, and if $\lambda \neq 0$ is any zero of the function $\mathcal{E}(x) - 1$, and m any integer, then*

$$\mathcal{E}(x + \bar{m}\lambda) = \mathcal{E}(x).$$

Furthermore one such zero λ always exists.

On utilizing the formulas of section 10, we can easily show that under the hypotheses of theorem 13.1, we also have

$$\sin\,(x + \bar{m}i\lambda) = \sin\,(x), \quad \cos\,(x + \bar{m}i\lambda) = \cos\,(x).$$

[7] The superscripts here denote ordinary differentiation.

IV. *The calculus of finite differences.*

14. Let (u) now be subject only to the conditions $u_0 = 0$, $u_1 = 1$, $u_n \neq 0$, $n \neq 0$. We shall denote by $\mathfrak{R}$ the ring of all polynomials in x with coefficients in the field of all complex numbers.

If

$$\phi = \phi(x) = \sum_{r=0}^{n} a_{n-r} x^r, \qquad a_0 \neq 0 \tag{14.1}$$

is any element of $\mathfrak{R}$ of degree n, we define the basic displacement symbol E by

$$\begin{aligned} E\phi(x) &= \phi(x+1) = \sum_{r=0}^{n} \sum_{s=0}^{r} a_{n-r}[r,s] x^{r-s} \\ E^{t+1}\phi(x) &= E(E^t\phi(x)), \quad E^0\phi(x) = \phi(x) \end{aligned} \tag{14.2}$$

where t is any positive integer.

It is obvious that E is a linear and distributive operator over $\mathfrak{R}$, and it may readily be shown that

$$E^t\phi(x) = \phi(x + \bar{t}). \tag{14.3}$$

If (u) is normal, formula (14.3) holds for all integral values of t.

15. The *basic difference operator* Δ is defined to be $E - 1$, where 1 stands for the identity operator over $\mathfrak{R}$. The following properties of Δ may be mentioned.

(i) Δ is linear and distributive over $\mathfrak{R}$, and converts an element of $\mathfrak{R}$ of degree n into one of degree $n-1$. Moreover E, Δ and D are commutative over $\mathfrak{R}$.

(ii) The only solutions of $\Delta\phi = 0$ lying in $\mathfrak{R}$ are $\phi =$ a constant.

(iii) $\Delta^t\phi(x) = \sum_{s=0}^{t} (-1)^s \binom{t}{s} \phi(x + \bar{s})$.

(iv) We have the operational identity over $\mathfrak{R}$

$$\Delta = \mathcal{E}(D) - 1 \tag{15.1}$$

where formally

$$\mathcal{E}(D) = \sum_{n=0}^{\infty} \frac{D^n}{[n]!}.$$

The last one of these properties is the only one requiring comment. If ϕ of formula (14.1) is operated on by D of section 7, then

$$\frac{D^s\phi(x)}{[s]!} = 0, \quad s > n; \qquad = \sum_{r=s}^{n} [r,s] a_{n-r} x^{r-s}, \quad s \leqq n.$$

Hence

$$\mathcal{E}(D)\phi(x) = \sum_{s=0}^{n} \sum_{r=s}^{n} [r,s] a_{n-r} x^{r-s} = \sum_{r=0}^{n} \sum_{s=0}^{r} [r,s] a_{n-r} x^{r-s} = E\phi(x)$$

by formula (14.2), so that (15.1) follows.

16. The *basic Bernouilli numbers* $B_0, B_1, \cdots, B_n, \cdots$ are defined by the recurrences

$$B_0 = 1; \quad (B+1)^n - B^n = 0, \; n > 1; \quad (B+1)^1 - B^1 = 1.$$

Here after expansion the exponents of B are to be degraded into suffices as in the usual theory [3].

The *basic Bernoulli polynomials* $B_n(z)$ may then be defined by

$$B_n(z) = (z+B)^n, \qquad (n = 0, 1, \cdots)$$

or non-symbolically,

$$B_n(z) = \sum_{r=0}^{n} [n, r] B_r z^{n-r}.$$

The following results [4] may be established precisely as in the ordinary theory.

$$(16.1) \qquad B_n(0) = B_n, \quad B_n(1) = B_n, \quad n \neq 1; \quad B_1(1) = B_1 + 1.$$

$$(16.2) \qquad B_n(x+y) = \sum_{r=0}^{n} [n, r] x^r B_{n-r}(y).$$

THEOREM 16.1. *If* $\phi'(x) = D\phi(x)$ *denotes the basic derivative of the polynomial* $\phi(x)$, *then a polynomial solution of the difference equation*

$$\Delta\Psi(x) = \phi'(x)$$

is given by

$$(16.3) \qquad \Psi(x) = \phi(x+B).$$

$$(16.31) \qquad \phi(x+B) = \sum_{r=0}^{n} \frac{\phi^{(r)}(x)}{[r]!} B_r = \sum_{r=0}^{n} \frac{\phi^{(r)}(0)}{[r]!} B_r(x).$$

$$(16.32) \qquad \Delta B_n(x) = [n] x^{n-1}.$$

THEOREM 16.2. *If the sequence* (u) *be chosen so that the series* (9.1) *for* $\mathcal{E}(x)$ *is convergent near* $x = 0$ *then for sufficiently small values of* $|t|$ *and* $|x|$

$$(16.4) \qquad \sum_{n=0}^{\infty} \frac{B_n t^n}{[n]!} = \frac{t}{\mathcal{E}(t) - 1}, \quad \sum_{n=0}^{\infty} \frac{B_n(x) t^n}{[n]!} = \frac{t\mathcal{E}(xt)}{\mathcal{E}(t) - 1}.$$

$$(16.5) \qquad \bar{1}^r + \bar{2}^r + \cdots + \overline{n-1}^r = \frac{B_{r+1}(\bar{n}) - B_{r+1}}{[r+1]},$$

if r *is a positive integer* $\geqq 1$.

To prove the last written formula for example, we observe by theorem 6.1 that (16.32) implies that $B_{r+1}(\overline{s+1}) - B_{r+1}(\bar{s}) = [r+1]\bar{s}^r$, s a positive integer. On summing this equation with respect to s from 0 to $n-1$, we obtain (16.5).

THEOREM 16.3. *$B_{2n} = 0$ $(n = 1, 2, 3, \cdots)$ when and only when*

$$B_n(1-z) = (-1)^n B_n(z), \quad (n = 2, 3, \cdots).$$

If moreover the series (9.1) *for $\mathcal{E}(x)$ converges for some $x \neq 0$, then*

$$B_{2n} = 0, \quad n \leqq 1$$

when and only when (u) is normal.

The equivalences stated follow immediately from formulas (16.1), (16.2) and (16.4).

We plan to give elsewhere a detailed treatment of the basic analogues for the numbers of Euler, Genocchi, Lucas and Stirling and their associated polynomials and difference operators.

REFERENCES.

1. F. H. Jackson, *American Journal of Mathematics*, vol. 32 (1910), pp. 305-314; *Messenger of Mathematics*, vol. 39 (1910), pp. 26-28, vol. 38 (1909), pp. 57-61, 62-64; *Proceedings of the Edinburg Mathematical Society*, vol. 22 (1904), pp. 28-39.

2. L. Euler, *Introductio in Analysin Infinitorum* (1748), chapter VII; Netto, *Combinatorik*, 2d. ed. (1927), p. 143.

3. D. H. Lehmer, *Annals of Mathematics* (2), vol. 36, no. 3, July (1935), p. 639 and references on p. 637.

4. Norlund, *Differenzenrechnung*, Berlin (1924), chapter II.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

ON THE TRANSFORMATION AND CONVERGENCE OF CONTINUED FRACTIONS.

By WALTER LEIGHTON and H. S. WALL.

1. *Introduction.* Convergence criteria for the continued fraction

$$y_0 + 1/y_1 + 1/y_2 + \cdots, \tag{1}$$

in which the y_n are complex numbers have been the object of study of many writers. E. B. Van Vleck [1] proved that if $y_1, y_3, y_5, \cdots$ are not all zero and if there exists a number $\epsilon > 0$ such that

$$-\pi/2 + \epsilon \leqq \arg y_n \leqq \pi/2 - \epsilon, \qquad (n = 1, 2, 3, \cdots)$$

a necessary and sufficient condition for the convergence of the continued fraction (1) is that the series $\Sigma \mid y_n \mid$ diverge. Pringsheim [2] showed that the conditions

$$\mid 1/y_{n-1}y_n \mid \leqq 1/4, \qquad (n = 2, 3, 4, \cdots) \tag{2}$$

are sufficient for the convergence of (1), with $\arg y_n$ unrestricted. Szász [3] pointed out that 1/4 is the largest number that can be used in (2).

When $\arg y_n$ is not restricted and the numbers $1/(y_n y_{n-1})$ do not satisfy the inequalities (2), the question of convergence of the continued fraction remains in a large measure unanswered.

In § 3 of this paper we have approached this problem by considering a new and very general transformation of a continued fraction. This approach has led to convergence theorems which are new and of a character fundamentally different from previously existing theorems. Some of these results are:

The continued fraction (1) *converges if*

$$(3)\quad 0 < \mid 1/y_{2n}y_{2n+1} \mid \leqq 1/4, \quad \mid 1/y_{2n-1}y_{2n} \mid \geqq 25/4, \quad (n = 1, 2, 3, \cdots);$$

or if

$$(4)\quad G_n \geqq \mid 1/y_{2n}y_{2n+1} \mid \geqq 25/4, \quad \mid 1/y_{2n-1}y_{2n} \mid \geqq 4(1 + G_n)(1 + G_{n-1}), \quad (n = 1, 2, 3, \cdots),$$

[1] E. B. Van Vleck, "On the convergence of continued fractions with complex elements," *Transactions of the American Mathematical Society*, vol. 2 (1901).

[2] O. Perron, *Die Lehre von den Kettenbrüchen*, 1 st. edition, p. 254.

[3] O. Szász, "Über die Erhaltung der Konvergenz Kettenbrüche bei independenter Veränderlichkeit aller ihrer Elemente," *Crelle*, Bd. 147 (1916), pp. 132-160.

where $G_0 = 0$ and $G_1, G_2, G_3, \cdots$ is any sequence of positive numbers.

The following theorem is of a different type.

Let $h_1, h_2, h_3, \cdots$ be any constants not zero such that for some integer $k \geq 0$

$$\lim_{n=\infty} h_{2n}h_{2n+2k-1} = 0, \qquad \lim_{n=\infty} h_{2n}h_{2n+2k+1} = \infty.$$

Then corresponding to every finite region R in the complex z-plane at a positive distance from the origin, there exists a positive integer N such that if $n \geq N$, the continued fraction

$$h_{2n} + z/h_{2n+1} + z/h_{2n+2} + \cdots$$

converges uniformly over R.

Still another kind of theorem is the following.

Let $q_1, q_2, q_3, \cdots$ be constants not zero. Let B represent an increasing sequence of indices $n_1 < n_2 < n_3 < \cdots$ in which $n_i - n_{i-1} \geq 2$, $i = 2, 3, 4, \cdots$, and let A represent the sequence of indices not in B. If B is finite and $\lim q_n = 0$, or if

$$\lim_{B} q_n = \infty, \qquad \lim_{A} q_n = 0,$$

the continued fraction

$$1 + q_n z/1 + q_{n+1} z/1 + \cdots$$

converges uniformly over a region R of the type described above, if n is taken sufficiently large.

2. *An important example.* In this section we indicate the complexity of the general convergence problem by exhibiting a convergent continued fraction (1) for which the numbers

$$x_n = 1/y_{n-1}y_n$$

are everywhere dense in the complex plane.

Let $a_1, a_2, a_3, \cdots$ be an infinite sequence of complex numbers not zero and set

$$A_n = a_1 + a_2 + \cdots + a_n.$$

Form the continued fraction (1) in which

$$(5) \qquad y_1 = \frac{2}{a_1}, \quad y_{2n} = -a_n, \quad y_{2n+1} = \frac{2(a_n + a_{n+1})}{a_n a_{n+1}}, \quad (n = 1, 2, 3, \cdots).$$

Its $2n$-th and $(2n+1)$-th convergents are the quantities A_n and $A_n + (a_{n+1}/2)$, respectively. Hence the continued fraction converges if and only if the series Σa_n converges.

Now let S denote any countable set of complex numbers $s_1, s_2, s_3, \cdots$ not containing 0 or $-1/2$. It is clear that given any positive number M and a number r such that $0 < r < 1$, one can find a sequence of numbers $k_1, k_2, k_3, \cdots$ for which the conditions

$$\left| \frac{2s_n}{k_n} \right| < M, \quad \left| \frac{1+2s_n}{k_n} \right| < M, \quad k_n r^2 + 2k_n r^2 s_{n+1} - 2s_n k_{n+1} \neq 0, \qquad (n = 1, 2, 3, \cdots),$$

are satisfied.

If we take

$$a_{2n} = \frac{-2s_n r^{2n}}{k_n}, \quad a_{2n-1} = \frac{(1+2s_n) r^{2n}}{k_n}, \qquad (n = 1, 2, 3, \cdots)$$

the series Σa_n and the continued fraction (1), with the numbers y_n given by (5), will converge. Moreover,

$$x_{4n+3} = s_{n+1}, \qquad (n = 0, 1, 2, \cdots).$$

Thus *the set S which may be taken everywhere dense in the complex plane, is contained in the set of elements $x_1, x_2, x_3, \cdots$ of a convergent continued fraction.*

From this example one concludes that *the convergence of the continued fraction* (1) *does not in general depend upon the magnitude of the $|x_n|$, but upon the relative magnitude of the numbers in the ordered sequence*

$$|x_2|, \quad |x_3|, \quad |x_4|, \cdots.$$

This principle would lead one to expect to find criteria such as those mentioned at the end of § 1, in which precisely this "relative magnitude" plays the dominant rôle.

3. *A transformation of continued fractions, and an extension of the Pringsheim criteria.* Let

$$(6) \qquad \xi \equiv (x_1, x_2, \cdots; y_0, y_1, y_2, \cdots) = y_0 + x_1/y_1 + x_2/y_2 + \cdots \qquad (x_n \neq 0)$$

and

$$(7) \qquad \eta \equiv (a_1, a_2, \cdots; b_0, b_1, b_2, \cdots) = b_0 + a_1/b_1 + a_2/b_2 + \cdots \qquad (a_n \neq 0)$$

be any two infinite continued fractions. Denote their n-th convergents by X_n/Y_n and A_n/B_n, respectively, where the quantities X_n, Y_n, A_n, B_n are given by the usual recursion relations. As is well-known,

$$\begin{vmatrix} X_n & X_{n+1} \\ Y_n & Y_{n+1} \end{vmatrix} = (-1)^{n+1} x_1 x_2 \cdots x_{n+1},$$

and hence, one can always determine numbers α_n, β_n such that

$$\begin{aligned} A_n &= \alpha_n X_n + \beta_n X_{n+1}, \\ B_n &= \alpha_n Y_n + \beta_n Y_{n+1}, \end{aligned} \qquad (n = 0, 1, 2, \cdots). \tag{8}$$

Conversely, corresponding to any continued fraction (6), let numbers α_n, β_n be any set for which

$$D_n \equiv \alpha_{n-1}(\alpha_n + \beta_n y_{n+1}) - \beta_n \beta_{n-1} x_{n+1} \neq 0, \quad n = 0, 1, 2, \cdots; \quad \alpha_1 = 1, \beta_{-1} = 0.$$

If then we put

$$E_n \equiv \alpha_n[\alpha_{n-2} y_n - \beta_{n-2} x_n] + \beta_n[\alpha_{n-2}(y_n y_{n+1} + x_{n+1}) - \beta_{n-2} x_n y_{n+1}],$$
$$n = 1, 2, 3, \cdots; \quad \alpha_{-1} = 1, \beta_{-1} = 0,$$

the continued fraction (7) with

$$\begin{aligned} b_0 &= y_0 + \beta_0 x_1 / D_0, \quad a_1 = x_1 D_1 / D_0^2, \\ a_n &= x_n D_n / D_{n-1}, \qquad (n = 2, 3, 4, \cdots), \\ b_n &= E_n / D_{n-1}, \qquad (n = 1, 2, 3, \cdots), \end{aligned} \tag{9}$$

will have the property that

$$\begin{aligned} D_0 A_n &= \alpha_n X_n + \beta_n X_{n+1}, \\ D_0 B_n &= \alpha_n Y_n + \beta_n Y_{n+1}, \end{aligned} \qquad (n = 0, 1, 2, \cdots). \tag{10}$$

Equations (10) are readily established by mathematical induction.

We shall consider the equations (9) as constituting a transformation T of ξ into η, and write

$$(a_1, a_2, \cdots; b_0, b, \cdots) = T(x_1, x_2, \cdots; y_0, y_1, \cdots)$$

or

$$\eta = T\xi.$$

The transformation is determined by the α_n, β_n, so that we shall write $T = T[\alpha_n; \beta_n]$. This notation will simplify the later discussion.

The use of special cases of this transformation such as the so-called "equivalent transformation" where

$$\alpha_0 = 1, \quad \alpha_n = c_1 c_2 \cdots c_n \qquad (c_i \neq 0); \quad \beta_n = 0,$$

is well-known.[4] We shall use the more general transformation in a comparable manner.

In order to apply the transformation T for obtaining convergence criteria for continued fractions it is important to establish conditions under which the convergence of $\eta = T\xi$ implies the convergence of ξ. For that purpose we employ the identity

$$\frac{A_n}{B_n} - \frac{X_n}{Y_n} \equiv \left(\frac{A_{n-1}}{B_{n-1}} - \frac{A_n}{B_n}\right)\left(\frac{\alpha_{n-1}B_n}{\beta_n x_{n+1}B_{n-1}} - 1\right)^{-1}. \tag{11}$$

We have the following result.

THEOREM 1. *If $\eta = T\xi$ converges, ξ converges to the same limit as η, provided, there exists a positive integer N and a positive constant c such that*

$$| \alpha_{n-1}B_n/\beta_n x_{n+1}B_{n-1} - 1 | \geqq C,$$

if $n > N$.

The proof is an easy consequence of (11).

If the elements a_n, b_n of η satisfy the inequalities

$$| b_n | \geqq | a_n | + 1, \qquad (n = 1, 2, 3, \cdots) \tag{12}$$

it may be shown [5] that η converges and the value of $\eta - b_0$ is numerically ≤ 1, and is < 1 if the inequality sign holds at least once in (12). Moreover, $| B_n/B_{n-1} | \geq 1$; hence, on referring to Theorem 1 we have the following theorem.

THEOREM 2. *Let $T[\alpha_n; \beta_n]$ be a transformation which carries the infinite continued fraction $\xi = (x_1, x_2, \cdots; y_0, y_1, y_2, \cdots)$, $(x_n \neq 0)$, into another infinite continued fraction $\eta = (a_1, a_2, \cdots; b_0, b_1, b_2, \cdots)$, $(a_n \neq 0)$, whose elements a_n, b_n satisfy (12). If there is a positive integer N and a number $s > 1$ such that for $\alpha_{n-1}\beta_n \neq 0$ and $n > N$*

$$| \alpha_{n-1}/\beta_n x_{n+1} | \geqq s, \tag{13}$$

then ξ and η converge to a common limit. The numerical value of $\xi - b_0$ is ≤ 1 and is < 1 if the inequality sign holds at least once in (12).

The foregoing theorem may be extended as follows. Let $c_1, c_2, c_3, \cdots$ be a sequence of constants none of which is zero. The continued fraction [6] $\eta' = (c_1a_1, c_1c_2a_2, c_2c_3a_3, \cdots; b_0, c_1b_1, c_2b_2, \cdots)$ converges if and only if η

[4] Perron, *loc. cit.*, p. 196.

[5] Perron, *ibid.*, pp. 254-255.

[6] Perron, *ibid.*, p. 196.

converges, and to the same limit. Let A'_n/B'_n be the n-th convergent of $(c_1c_2a_2, c_2c_3a_3, \cdots; 0, c_2b_2, c_3b_3, \cdots)$. If the latter converges to a value different from $-c_1b_1$, η' converges. If then we assume that

$$|c_1b_1| \geqq 1, \quad |c_nb_n| \geqq |c_{n-1}c_na_n| + 1, \qquad (n = 2, 3, 4, \cdots), \tag{14}$$

where inequality holds at least once, it must follow that the limit

$$\lim_{n=\infty} [c_1b_1 + A'_n/B'_n] \tag{15}$$

exists and is $\neq 0$. If we set

$$v_n = c_1b_1 + A'_n/B'_n$$

and denote the limit (15) by v, we observe that

$$\eta = b_0 + (c_1a_1/v),$$

$$\left| \frac{\alpha_{n-1}B_n}{\beta_nx_{n+1}B_{n-1}} \right| = \left| \frac{\alpha_{n-1}B'_{n-1}v_{n-1}}{c_n\beta_nx_{n+1}B'_{n-2}v_{n-2}} \right|,$$

and that $|B'_{n-1}/B'_{n-2}| \geq 1$. We have thus proved the following theorem.

THEOREM 3. *Let $T[\alpha_n; \beta_n]$ be a transformation which carries the infinite continued fraction $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$, $(x_n \neq 0)$, into another infinite continued fraction $\eta = (a_1, a_2, \cdots; b_0, b_1, \cdots)$ whose elements satisfy conditions* (14) *with actual inequality holding at least once for some set of numbers $c_1, c_2, c_3, \cdots$, $(c_i \neq 0)$. If there exists a positive integer N and a number $s > 1$ such that*

$$\left| \frac{\alpha_{n+1}}{c_n\beta_nx_{n+1}} \right| \geqq s \tag{16}$$

for all values of $n > N$ for which $\alpha_{n-1}\beta_n \neq 0$, then ξ and η converge to a common limit.

If, in particular, we set $c_n = p_n/b_n$ where p_n is real and positive we have the following result.

THEOREM 4. *The inequalities* (14) *and* (16) *of Theorem* 3 *may be replaced by the following inequalities:*

$$p_1 \geqq 1, \left| \frac{a_n}{b_{n-1}b_n} \right| \leqq \frac{p_n - 1}{p_np_{n-1}}, \quad (n = 2, 3, 4, \cdots); \tag{14'}$$

$$\left| \frac{\alpha_{n-1}b_n}{\beta_nx_{n+1}} \right| \geqq sp_n, \tag{16'}$$

respectively, where $p_1, p_2, p_3, \cdots$ *are real numbers.*

The special case $p_n = 2$ is important. In this case we readily obtain the following theorem.[7]

THEOREM 5. *If the* x_n, y_n *in* $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$, $(x_n \neq 0)$, *are functions of any variables, then* ξ *converges uniformly over the region determined by the inequalities*

$$\left|\frac{a_1}{b_1}\right| \leqq C, \quad \left|\frac{a_n}{b_{n-1}b_n}\right| \leqq 1/4, \quad (n = 2, 3, 4, \cdots),$$
$$\left|\frac{\alpha_{n-1}b_n}{\beta_n x_{n+1}}\right| \geqq 2s, \quad n > N,$$

where $C, N,$ *and* s *are constants and* $s > 1$.

4. *The transformation* T_1. Let

$$\alpha_{2n} = 1 - g_n y_{2n+1}, \quad \alpha_{2n+1} = -y_{2n+2}/g_n, \quad \beta_{2n} = g_n, \quad \beta_{2n+1} = 1/g_n, \quad (n = 0, 1, 2, \cdots),$$

where

$$g_n = g_0 - (y_2 + y_4 + \cdots + y_{2n}), \qquad (n = 1, 2, 3, \cdots),$$

and g_0 is any number $\neq 0$ so chosen that $g_n \neq 0$, $n = 1, 2, 3, \cdots$. If $x_n = 1$, the transformation T_1 is as follows:

$$(T_1) \qquad \begin{aligned} &b_0 = y_0 + g_0, \ a_1 = -1, \ a_n = 1, \ n > 1; \\ &b_1 = 1/g_0, \ b_{2n} = y_{2n-1}g^2_{n-1}, \qquad (n = 1, 2, 3, \cdots). \\ &b_{2n+1} = y_{2n}/g_n g_{n-1}, \end{aligned}$$

Here

$$(17) \qquad A_{2n+1}/B_{2n+1} = X_{2n}/Y_{2n}, \qquad (n = 0, 1, 2, \cdots),$$

so that Theorem 5 in this case takes the following form.

THEOREM 6. *Let* $y_0, y_1, y_2, \cdots$ *be functions of any variables, and let* $g_0 \neq 0$ *be so chosen that* $g_n = g_0 - (y_2 + y_4 + \cdots + y_{2n}) \neq 0$, $n = 1, 2, 3, \cdots$. *Set*

$$t_1 = -g_0, \ t_2 = \frac{1}{y_1 g_0}, \ t_{2n} = \frac{g_{n-2}}{y_{2n-1}y_{2n-2}g_{n-1}}, \ n > 1, \ t_{2n+1} = \frac{g_n}{y_{2n}y_{2n-1}g_{n-1}}, \ n > 0;$$
$$k_n = \frac{1}{g_n y_{2n+1}}, \ n > 0.$$

[7] Perron, *ibid.*, pp. 260-262.

Then the continued fraction $(1, 1, 1, \cdots; y_0, y_1, y_2, \cdots)$ *converges uniformly over the region determined by the inequalities:*

$$\begin{aligned} &|t_1| < C, \quad |t_n| \leqq 1/4, && (n = 2, 3, 4, \cdots) \\ &|t_{2n+2}/(k_n - 1)| \leqq 1/2s, \quad n > N, \end{aligned}$$

where C, N, *and* s *are constants and* $s > 1$.

As an application we have this theorem:

THEOREM 7. *Let* $h_0, h_1, h_2, \cdots$ *be real or complex numbers* $h_i \neq 0$, *and let* $q_n = q_0 - (h_2 + h_4 + \cdots + h_{2n})$, *where* $q_0 \neq 0$ *is so chosen that* $q_n \neq 0$, $n = 1, 2, 3, \cdots$. *Put*

$$w_{2n} = \frac{q_{n-2}}{q_{n-1}h_{2n-1}h_{2n-2}}, \ n > 1, \qquad w_{2n+1} = \frac{q_n}{q_{n-1}h_{2n}h_{2n-1}}, \ n > 0,$$

and suppose $\lim w_n = 0$. *Then the continued fraction*

$$H = h_0 + z/h_1 + z/h_2 + \cdots$$

has the following properties.

(a) *The sequence of even convergents represents a meromorphic function* $F(z)$ *and converges uniformly over every closed bounded region* R *containing no poles of* $F(z)$.

(b) *Let* $r_n = q_n h_{2n+1}$, *and let* R' *be an arbitrary closed bounded region containing none of the poles of* $F(z)$, *and no limit points of the sequence* $\{r_n\}$. *Then the sequence of odd convergents converges uniformly over* R' *to* $F(z)$.

This theorem is a notable extension of an earlier theroem [8] in which was required the absolute convergence of the series Σw_n.

Write H in the form

$$H' = (1, 1, 1, \cdots; h_0, h_1 z', h_2, h_3 z', \cdots), \qquad z' = 1/z.$$

Then $T_1 H'$ can be put in the form

$$\eta = (w_1 z, w_2 z, \cdots; h_0 + q_0, 1, 1, \cdots).$$

Since $\lim w_n = 0$, η represents a meromorphic function $F(z)$ and converges uniformly over R. By (12) the same is true of the sequence of even convergents of H. This proves (a).

[8] Wall, *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 946-952.

To prove (b), apply Theorem 6 to the continued fraction

$$\xi_{\mu,\nu} = h_{2\mu} + \frac{1}{\nu z' h_{2\mu+1}} + \frac{1}{h_{2\mu+2}} + \frac{1}{z' h_{2\mu+3}} + \cdots, \quad z' = 1/z,$$

taking $g_0 = q_0 - (h_2 + h_4 + \cdots + h_{2\mu})$. Then

$$t_2 = \frac{z}{\nu q_\mu h_{2\mu+1}}, \quad t_\nu = z w_{2\mu+\nu}, \qquad (\nu = 3, 4, 5, \cdots).$$

If we now assign to μ and then to ν sufficiently large values, we will have

$$| t_\nu | \leqq 1/4, \qquad (\nu = 2, 3, 4, \cdots),$$

over a circle K with center $z = 0$ and a sufficiently large radius to include the region R'. Also, over R':

$$\left| \frac{t_{2n+2}}{k_n - 1} \right| = \left| \frac{\dfrac{z w_{2\mu+2n+2}}{z}}{\dfrac{r_{n+\mu}}{r_{n+\mu}} - 1} \right| \leqq \frac{1}{2s},$$

if $n >$ some N, and s is a constant > 1. Hence, by Theorem 6, $\xi_{\mu,\nu}$ converges uniformly over R' to a function $f(z)$ which is analytic over K, and is non-rational. Consequently,

$$\frac{1}{h_{2\mu+2}} + \frac{1}{h_{2\mu+3} z'} + \frac{1}{h_{2\mu+4}} + \cdots$$

converges over R' to the value

$$f_1(z) = \frac{1}{f(z) - h_{2\mu}} - \nu z' h_{2\mu+1},$$

and H converges to

$$F(z) = \frac{X_{2\mu+1} + f_1(z) X_{2\mu}}{Y_{2\mu+1} + f_1(z) Y_{2\mu}},$$

which is analytic, except possibly for poles, inasmuch as $f_1(z)$ is non-rational. The convergence is clearly uniform over R'.

THEOREM 8. *Let* $(1, 1, 1, \cdots; 0, y_1, y_2, \cdots) = \xi$ *be a continued fraction with arbitrary partial denominators* $y_1, y_2, y_3, \cdots$; *and suppose the series*

$$\sum_{n=1}^{\infty} y_{2n+1} (y_2 + y_4 + \cdots + y_{2n})^2, \tag{18}$$

$$\sum_{n=m}^{\infty} \left[\frac{1}{y_2 + y_4 + \cdots + y_{2n-2}} - \frac{1}{y_2 + y_4 + \cdots + y_{2n}} \right], \quad m > 1,$$

converge absolutely, and that

$$\lim_{n=\infty} | y_2 + y_4 + \cdots + y_{2n} | = \infty. \tag{19}$$

Then there exist two numbers F, G, where $|F|+|G|\neq 0$, such that

$$\lim_{n=\infty}\frac{X_{2n}}{y_2+y_4+\cdots+y_{2n}}=\lim_{n=\infty}X_{2n+1}=F,$$
$$\lim_{n=\infty}\frac{Y_{2n}}{y_2+y_4+\cdots+y_{2n}}=\lim_{n=\infty}Y_{2n+1}=G. \tag{20}$$

If $y_1, y_2, y_3, \cdots$ are analytic functions of z over a region R in which (18) *converge absolutely and uniformly, then F, G are analytic functions over R,* and *ξ converges to F/G, wherever $G\neq 0$.*

If (18) *converge absolutely and* (19) *fails to hold, then ξ diverges.*

To prove the theorem, apply the transformation T_1 to ξ, and obtain

$$\frac{X_{2n}}{g_n}=A_{2n+1}, \qquad X_{2n+1}=\frac{A_{2n}}{g_n}-(1-g_ny_{2n+1})A_{2n+1}, \tag{21}$$

with like expressions for the Y_n. Now from the absolute convergence of (18) it follows that the series

$$\sum\frac{y_{2n}}{g_{n-1}g_n}, \qquad \sum y_{2n+1}{g_n}^2 \tag{22}$$

converge absolutely, and hence by a theorem of von Koch:[9]

$$\begin{aligned}&\lim A_{2n+1}=F, &&\lim A_{2n}=F_0,\\ &\lim B_{2n+1}=G, &&\lim B_{2n}=G_0,\end{aligned} \tag{23}$$

where $FG_0-F_0G=1$, so that $|F|+|G|\neq 0$. If (18), and hence (22), converge absolutely and uniformly over R, then F, G, F_0, G_0 are analytic functions over R. The equations (20) now follow at once from (21), (23), and the hypothesis.

It is obvious that when (18) converge absolutely and (19) fails to hold, then $\Sigma|y_n|$ converges, and hence ξ diverges.

Remark. The second series (18) evidently converges when (19) holds. If we could remove the requirement of absolute convergence of this series, this theorem would include the following theorem of Hamburger:[10] *Let $(1, 1, 1, \cdots; 0, a_1z, a_2, a_3z, \cdots)$ be a continued fraction in which the a_n are real and $\neq 0$, and $a_{2n+1}>0$. Then, if the series $\Sigma a_{2n+1}(a_2+a_4+\cdots+a_{2n})^2$ converges and $\lim|a_2+a_4+\cdots+a_{2n}|=\infty$, the continued fraction converges to a function which is analytic except for poles.*

---

[9] Von Koch, *Bulletin de la Société Mathématique de France*, vol. 23 (1895).

[10] H. Hamburger, *Mathematische Annalen*, vol. 82, pp. 120-164, 168-187.

5. *Transformations which interchange two convergents.* We proceed to obtain a transformation $T = W_m$ which interchanges the m-th and $(m+1)$-th convergents of $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$. The interchange of any two convergents can be effected by repeated application of W_m with properly chosen values of m. We find that the formulas for the transformation W_m are as follows:

$$(W_0)\quad \begin{cases} b_0 = y_0 + \dfrac{x_1}{y_1},\ a_1 = -\dfrac{x_1}{y_1^2},\ a_2 = y_2,\ a_3 = -\dfrac{x_2 x_3}{y_2}, \\ b_1 = \dfrac{1}{y_1},\ b_2 = x_2,\ b_3 = \dfrac{y_2 y_3 + x_3}{y_2};\ a_n = x_n,\ b_n = y_n, \\ \qquad\qquad (n = 4, 5, 6, \cdots), \end{cases}$$

$$(W_m)\quad \begin{cases} a_m = x_m y_{m+1},\ a_{m+1} = -x_{m+1}/y_{m+1},\ a_{m+2} = y_{m+2}, \\ a_{m+3} = -x_{m+2}x_{m+3}/y_{m+2},\ b_m = y_m y_{m+1} + x_{m+1}, \\ b_{m+1} = 1/y_{m+1},\ b_{m+2} = x_{m+2},\ b_{m+3} = \dfrac{y_{m+2}y_{m+3} + x_{m+3}}{y_{m+2}}, \\ a_n = x_n, \qquad (n \neq m,\ m+1,\ m+2,\ m+3) \\ b_n = y_n \qquad (m = 1, 2, 3, \cdots). \end{cases}$$

This transformation is valid if $y_{m+1}y_{m+2} \neq 0$. If $b_n, y_n \neq 0$, $n = 1, 2, 3, \cdots$, and we set

$$\begin{aligned} t_1 &= x_1/y_1, \quad t_n = x_n/y_{n-1}y_n, \\ t'_1 &= a_1/b_1, \quad t'_n = a_n/b_{n-1}b_n, \end{aligned} \qquad (n = 2, 3, 4, \cdots),$$

the transformation becomes:

$$(W_0)\quad \begin{cases} b_0 = y_0 + t_1,\ t'_1 = -t_1,\ t'_2 = 1/t_2,\ t'_3 = \dfrac{-t_3}{1+t_3},\ t'_4 = \dfrac{t_4}{1+t_3}, \\ t'_n = t_n, \qquad (n = 5, 6, 7, \cdots); \end{cases}$$

$$(W_m)\quad \begin{cases} t'_m = \dfrac{t_m}{1+t_{m+1}},\ t'_{m+1} = -\dfrac{t_{m+1}}{1+t_{m+1}},\ t'_{m+2} = \dfrac{1}{t_{m+2}}, \\ t'_{m+3} = \dfrac{-t_{m+3}}{1+t_{m+3}},\ t'_{m+4} = \dfrac{t_{m+4}}{1+t_{m+3}}, \\ t'_n = t_n \quad (n = 1, 2, 3, \cdots, m-1,\ m+5,\ m+6, \cdots), \\ \qquad\qquad (m = 1, 2, 3, \cdots). \end{cases}$$

Let T_2 denote the transformation which results by applying in succession $W_0, W_2, W_4, \cdots$. Then the formulas for T_2 are:

$$(T_2)\quad \begin{cases} b_0 = y_0 + t_1,\ t'_1 = -t_1,\ t'_2 = \dfrac{1+t_3}{t_2}, \\ t'_{2n} = \dfrac{(1+t_{2n-1})(1+t_{2n+1})}{t_{2n}}, \qquad (n = 2, 3, 4, \cdots), \\ t'_{2n+1} = t_{2n+1}, \qquad (n = 1, 2, 3, \cdots). \end{cases}$$

If we apply in succession $W_1, W_4, W_7, W_{10}, \cdots$, we obtain:

$$(T_3)\quad \begin{cases} b_0 = y_0,\ t'_1 = \dfrac{t_1}{1+t_2},\ t'_2 = \dfrac{-t_2}{1+t_2}, \\ t'_{3n} = \dfrac{1}{t_{3n}},\ t'_{3n+1} = \dfrac{-t_{3n+1}}{1+t_{3n+1}+t_{3n+2}},\ t'_{3n+2} = \dfrac{-t_{3n+2}}{1+t_{3n+1}+t_{3n+2}}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (n=1,2,\cdots). \end{cases}$$

These transformations furnish immediate extension of Pringsheim's criteria. For example, from W_0 we obtain this result:

THEOREM 9. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, 1, \cdots)$ $(t_n \neq 0)$ *converges if*

$$|t_2| \geqq 4, \quad |t_3| \leqq 1/5, \quad |t_4| \leqq 1/5, \quad |t_n| \leqq 1/4, \qquad (n=5,6,7,\cdots).$$

From T_2 we obtain the theorem:

THEOREM 10. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, 1, \cdots)$ $(t_n \neq 0)$ *converges if there exist real numbers* $p_1, p_2, p_3, \cdots, p_i > 1$, *such that*

$$|t_{2n+1}| \leqq \frac{p_{2n+1}-1}{p_{2n}p_{2n+1}}, \qquad |t_{2n}| \geqq M_{2n}, \qquad (n=1,2,3,\cdots)$$

where

$$M_2 = \frac{p_1(p_2p_3+p_3-1)}{p_3(p_2-1)}, \quad M_{2n} = \frac{(p_{2n-2}p_{2n-1}+p_{2n-1}-1)(p_{2n}p_{2n+1}+p_{2n+1}-1)}{(p_{2n}-1)(p_{2n-2}p_{2n+1})} \qquad (n>1).$$

If $p_n = 2$ this gives the first theorem mentioned at the end of the introduction.

From T_3 we have:

THEOREM 11. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, \cdots)$ $(t_n \neq 0)$ *converges if*

$$|t_2| \leqq 1/5, \quad |t_{3n}| \geqq 4, \quad |t_{3n+1}| \leqq 1/6, \quad |t_{3n+2}| \leqq 1/6, \qquad (n=1,2,3,\cdots).$$

We shall now apply the transformation W_m to prove the following theorem mentioned at the end of § 1.

THEOREM 12. *Let* $q_1, q_2, q_3, \cdots$ *be constants* $\neq 0$, *and let*

$$\lim_{\nu=\infty} q_{n'_\nu} = \infty, \qquad \lim_{\nu=\infty} q_{n''_\nu} = 0,$$

where $B = (n_1' < n_2' < \cdots)$ *and* $A = (n_1'' < n_2'' < \cdots)$ *are mutually exclusive sets whose sum is the set of natural numbers* $1, 2, 3, \cdots$; *and suppose that if* n *is in* B, *then* $n - 1$ *and* $n + 1$ *are in* A. *Then, corresponding to every bounded region* R *of the complex* z*-plane whose distance from the origin is* > 0, *there exists an index* N *such that if* $n \geqq N$ *the continued fraction*

$$1 + \frac{q_n z}{1} + \frac{q_{n+1} z}{1} + \cdots \tag{24}$$

converges uniformly over R.

Proof. Choose N sufficiently large to insure that, over R,

$$| q_\nu z | \leqq \delta, \ \nu \geqq N, \ \nu \text{ in } A; \quad \left| \frac{1}{q_\nu z} \right| \leqq \delta, \ \nu \geqq N, \ \nu \text{ in } B,$$

where δ is a properly chosen small positive number. Let $n'_\mu, n'_{\mu+1} \cdots$ be the numbers in B which are $> n \geqq N$, and apply to (24) in succession the transformations

$$W_{n'_i - 2}, \qquad (i = \mu, \mu + 1, \mu + 2, \cdots).$$

Then in the resulting continued fraction, $(t'_1, t'_2, t'_3, \cdots; 1, 1, 1, \cdots)$, we shall have, over R,

$$| t'_n | \leqq 1/4, \qquad (n = 2, 3, 4, \cdots)$$

and hence, (24) converges uniformly over R.

6. *The transformation* T_4. This transformation replaces the $2n$-th convergent by the $(2n + 1)$-th, $(n = 0, 1, 2, \cdots)$. The formulas are found to be

$$(T_4) \qquad \begin{gathered} b_0 = y_0 + \frac{t_1}{1 + t_2}, \qquad t'_1 = \frac{t_1 t_2}{1 + t_2}, \\ t'_{2n} = \frac{(1 + t_{2n})(1 + t_{2n+2})}{t_{2n+1}}, \qquad t'_{2n+1} = t_{2n+2}, \qquad (n = 1, 2, 3, \cdots). \end{gathered}$$

If we follow the transformation T_2 by T_4 we shall obtain

$$b_0 = y_0 - \frac{t_1 t_2}{1 + t_2 + t_3}, \quad t'_1 = \frac{-t_1(1 + t_3)}{1 + t_2 + t_3},$$

$$T'_2 = \frac{\left(1 + \dfrac{1 + t_3}{t_2}\right)\left(1 + \dfrac{(1 + t_3)(1 + t_5)}{t_4}\right)}{t_3},$$

$$(T_4T_2)\quad t'_{2n}=\frac{\left(1+\dfrac{(1+t_{2n-1})(1+t_{2n+1})}{t_{2n}}\right)\left(1+\dfrac{(1+t_{2n+1})(1+t_{2n+3})}{t_{2n+2}}\right)}{t_{2n+1}},\qquad (n=2,3,4,\cdots),$$

$$t'_{2n+1}=\frac{(1+t_{2n+1})(1+t_{2n+3})}{t_{2n+2}},\qquad (n=1,2,3,\cdots).$$

From this, Pringsheim's criterion gives the following theorem.

THEOREM 13. *Let $p_1, p_2, p_3, \cdots$ be real and >1, and let $M_n=\dfrac{p_n-1}{p_np_{n-1}}$, $n>1$, $M_1>0$,*

$$\bar{M}_{2n}=\frac{(1+M_{2n-1})(1+M_{2n+1})}{M_{2n}},\qquad (n=1,2,3,\cdots).$$

Let $M'_{2n}\geqq\bar{M}_{2n}$, and set $\bar{M}'_{2n}=(1+M'_{2n})(1+M'_{2n-2})/M_{2n-1}$, $n=1,2,3,\cdots$ $M'_0=0$. Then the continued fraction $(x_1, x_2, \cdots; y_0, 1, 1, \cdots)$ converges if

$$(25)\qquad M'_{2n}\geqq|t_{2n+1}|\geqq\bar{M}_{2n},\qquad |t_{2n}|\geqq\bar{M}'_{2n},\qquad (n=1,2,3,\cdots).$$

If $p_n=2$, $M'_{2n}=M$, then (25) becomes

$$(25')\qquad M\geqq|t_{2n+1}|\geqq 25/4,\quad |t_{2n}|\geqq 4(1+M)^2,\qquad (n=1,2,3,\cdots).$$

If $p_n=2$ and $\lim M'_{2n}=\infty$, then $\lim|t_{2n}/t_{2n+1}|=\infty$.

By application of the transformations $T_2, T_4T_2, T_4^2T_2, \cdots$ we obtain

THEOREM 14. *Let $h_1, h_2, h_3, \cdots$ be any numbers $\neq 0$ such that for some $k\geqq 0$:*

$$(26)\qquad \lim_{n=\infty}h_{2n}h_{2n+2k-1}=0,\qquad \lim_{n=\infty}h_{2n}h_{2n+2k+1}=\infty.$$

Then, corresponding to every bounded region R of the complex z-plane whose distance from the origin is positive, there exists an index N such that if $n\geqq N$ the continued fraction

$$(27)\qquad h_{2n}+\frac{z}{h_{2n+1}}+\frac{z}{h_{2n+2}}+\cdots$$

converges uniformly over R.

To prove the theorem, let $q_1=1/h_1$, $q_n=1/h_{n-1}h_n$, $n>1$. Then (27) takes the form

$$h_{2n}\left[1+\frac{q_{2n+1}z}{1}+\frac{q_{2n+2}z}{1}+\cdots\right]\equiv h_{2n}\xi_n;$$

and $T_2\xi_n = (t'_1, t'_2, \cdots; b_0, 1, 1, \cdots)$, where

$$t'_2 = \frac{1 + q_{2n+1}z}{q_{2n}z}, \qquad t'_{2n} = \frac{(1 + q_{2n-1}z)(1 + q_{2n+1}z)}{q_{2n}z},$$

$$n > 1, \; t'_{2n+1} = q_{2n+1}z, \; n > 0.$$

Hence, if $\lim q_{2n+1} = \lim (1/h_{2n}h_{2n+1}) = 0$, $\lim q_{2n} = \lim (1/h_{2n}h_{2n+1}) = \infty$, it is clear that there exists an index N such that if $n \geqq N$, $| t'_\nu | \leqq 1/4$, $\nu = 2, 3, 4, \cdots$, over R, so that ξ_n, and therefore (27), converge uniformly over R. This proves the theorem for the case $k = 0$.

In like manner, when $k = 1, 2, \cdots$ the theorem may be proved by using the transformations $T_4T_2, T_4{}^2T_2, \cdots$, respectively. The proof for all k can be readily accomplished by mathematical induction.

Example. Let $(z, z, z, \cdots; 0, h_1, h_2, h_3, \cdots)$ be a Stieltjes continued fraction in which the h_n are real and positive. Then if (26) holds, Σh_ν diverges and hence the continued fraction converges, except along the whole or a part of the negative half of the real axis. Inasmuch as (27) is an analytic function of z over R, if n is sufficiently large, *it follows that our Stieltjes continued fraction represents a function whose only singularities in the region R are poles.* The same argument can be used if the h_n are real and $\neq 0$ and $h_{2n+1} > 0$. The conditions of Theorem 14 are met if

$$h_{2n} = r^{n^2}u_{2n}, \qquad h_{2n-1} = \frac{u_{2n-1}}{r^{(n-k-1/2)^2}}, \qquad (n = 1, 2, 3, \cdots),$$

where $0 < r < 1$, and $u_1, u_2, u_3, \cdots$ are any constants which are bounded and bounded away from 0.

HARVARD UNIVERSITY
AND
NORTHWESTERN UNIVERSITY.

THE FORM $2wx + xy + yz + zu + ux$.

By E. T. Bell.

1. In a recent paper [1] I gave a set of 25 arithmetical identities equivalent to all those of a certain sort between doubly periodic functions of the first and second kinds, all referring to the arithmetical form $xy + zw$. From the same identities another set concerning the form $wx + xy + yz$, or the like with numerical coefficients > 1, can be derived, and these in turn yield interesting results concerning numbers of representations in quinary quadratic forms. To illustrate one method of obtaining the new identities we shall prove the following.

THEOREM 1. *Let t, τ, w, x, y, z be integers > 0 such that, for m odd and constant,*

$$m = t\tau = wx + xy + 2yz, \tag{1}$$

and let $f(u)$ be any even function of u which is defined for integer values of u,

$$f(u) = f(-u),\ u \text{ integral}. \tag{2}$$

Then

$$\Sigma[(x+z)f(w-y) - zf(w+y)] = \Sigma\sum_{r=1}^{(\tau-1)/2}[rf(t) - tf(\tau+2-4r)], \tag{3}$$

the Σ on the left referring to all (w, x, y, z), that on the right to all (t, τ) satisfying (1).

Let $N[n = F; *]$ denote the number of representations of n in the form F, the integer variables in F being restricted by the conditions $*$. We shall prove from (3),

THEOREM 2.

$$\zeta_2(m) - (4m+1)\zeta_0(m) + 4\zeta_1(m) = 8N[m = 2wx + xy + yz + zu + ux;\ w, x, y, z > 0;\ u \geqq 0],$$

in which m is odd, and $\zeta_r(m)$ denotes the sum of the r-th powers of all the divisors of m (so that $\zeta_0(m)$ is the number of divisors).

This result is unusual in that it expresses the number of representations in an appropriately restricted indefinite form in five variables in terms of the

[1] *American Journal of Mathematics*, vol. 57 (1935), pp. 245-253.

divisors of the integer represented. There seems to be but one other similar result in the literature, stated in 1866 by Liouville (a proof of which will be published elsewhere).

2. In the formula (II), p. 249 of the paper cited we may choose $|\, y \,|\, f(x)$ for the function of x, y there, where $f(x)$ is as in (2). We get

$$(4)\quad \sum [(\tau_1 + \tau_2) f(t_1 - 2t_2) - |\, \tau_1 - \tau_2 \,|\, f(t_1 + 2t_2)] = 2 \sum \sum_{r=1}^{(\tau-1)/2} r f(t),$$

the sums referring to all $(t_1, \tau_1, t_2, \tau_2)$, (t, τ) from

$$(5)\quad m = t\tau = t_1\tau_1 + 2t_2\tau_2,$$

in which all letters denote integers > 0, m, τ_2 are odd (and hence also t_1, τ_1).

Considering the second term in (4) we distinguish the cases $\tau_1 > \tau_2$, $\tau_1 < \tau_2$, $\tau_1 = \tau_2$, and make the corresponding changes in (5):

$$(6)\quad \tau_1 > \tau_2;\ \tau_1 = \tau_2 + 2x,\ x > 0;\quad m = t_1(\tau_2 + 2x) + 2t_2\tau_2,\ t_1, \tau_2 \text{ odd};$$

$$(7)\quad \tau_2 > \tau_1;\ \tau_2 = \tau_1 + 2y,\ y > 0;\quad m = t_1\tau_1 + 2t_2(\tau_1 + 2y),\ t_1, \tau_1 \text{ odd};$$

$$(8)\quad \tau_1 = \tau_2;\ m = t\tau = \tau_1(t_1 + 2t_2).$$

Solving (8), we have $\tau_1 = \tau$, $t_1 + 2t_2 = t$. For (t, τ) a fixed pair of divisors of m there are the solutions

$$\{t_1, t_2\} = \{2j - 1, (t - 2j + 1)/2\}, \qquad (j = 1, \cdots, (t-1)/2).$$

Hence (8) contributes

$$(9)\quad 2 \sum \tau \sum_{j=1}^{(\tau-1)/2} f(t - 4j + 2)$$

to the left of (4), Σ referring to all (t, τ) in (5). We may interchange t, τ in (9). Thus (4) becomes

$$(10)\quad \begin{aligned} &\sum_{6} [(\tau_2 + x) f(t_1 - 2t_2) - x f(t_1 + 2t_2)] \\ &+ \sum_{7} [(\tau_1 + y) f(t_1 - 2t_2) - y f(t_1 + 2t_2)] \\ &= \sum \sum_{r=1}^{(\tau-1)/2} [r f(t) - t f(\tau + 2 - 4r)], \end{aligned}$$

where $\sum_6$, $\sum_7$ refer to the forms in (6), (7). Both forms are included in

$$m = w(\tau_2 + 2x) + v\tau_2,\ \tau_2 \text{ odd},\ w \not\equiv v \bmod 2.$$

Since m is odd, the conditions on τ_2, w, v are automatically satisfied and may be suppressed. Using the condition (2) we have

$$\sum_6 + \sum_7 = \sum [(y+x)f(w-v) - xf(w+v)],$$

the sum referring to all sets (v, y, w, x) of integers > 0 such that

$$m = vy + yw + 2wx.$$

The change of notation $(v, y, w, x) \to (w, x, y, z)$ then gives the result in Theorem 1 from (10).

3. To prove Theorem 2 take $f(x) = 1$ for all integer values of x in (3). A simple reduction gives

$$8\sum x = \zeta_2(m) - (4m+1)\zeta_0(m) + 4\zeta_1(m),$$

the sum extending over all solutions of (1). But

$$x = N[x = x_1 + x_2;\ x_1 > 0,\ x_2 \geqq 0].$$

Hence we get the result stated by the change of notation

$$(z, y, x_1, w, x_2) \to (w, x, y, z, u).$$

California Institute of Technology.

REDUCIBLE EXCEPTIONAL CURVES.

By PATRICK DU VAL.

The following research was undertaken and to a great extent carried out independently of that of Barber and Zariski on the same topic.[1] It may be regarded as a continuation, extension, and I hope a simplification, of the results of those authors. They only consider in much detail reducible curves transformable into the neighbourhoods of a sequence of points along a single branch; but I shall shew that very simple and elegant results hold for the neighbourhoods of a set of points grouped round each other in any way.

The essential idea for which I may refer the reader to the paper already mentioned is that each point of the set has two neighbourhoods, a total neighbourhood which is a curve (generally reducible) of virtual genus 0 and virtual grade -1, and a diminished neighbourhood which is an irreducible rational curve, of virtual grade generally < -1. The total neighbourhood of a point not in the neighbourhood of any other, i. e. of what we may call an original point, is what is subtracted from any linear system not having a base point at the point in question, by giving it a simple base point there; the diminished neighbourhood is what is left of this after those of its points which belong to the set in question have been transformed into curves. It is evident from this that the diminished neighbourhood of the point is equal to the total neighbourhood, less the total neighbourhoods of certain other points of the set. For a point which is not original we apply the definition by first transforming all its ancestor points into curves, when it becomes original.

It is evident that the points can be arranged in such an order that none is in a neighbourhood of any which follows it. This we shall call a standard order. It is unique if and only if the points are consecutive on a single branch.

We shall suppose the set of points to be n in number, and call them $O_1, \cdots, O_n$, in a standard order. The total and diminished neighbourhoods of O_a we shall call Q_a, L_a, respectively; the set of curves Q_a, arranged as a matrix of one row and n columns, will be denoted by $\boldsymbol{Q}$, the same set arranged as a matrix of one column and n rows by $\tilde{\boldsymbol{Q}}$; $\boldsymbol{L}$ and $\tilde{\boldsymbol{L}}$ will have similar meanings. (Throughout, clarendon type will indicate matrices, and the symbol $\tilde{}$ the transposition of a matrix, or interchange of rows and columns.) From what has been said it is clear that there is a relation

[1] S. F. Barber and Oscar Zariski, "Reducible exceptional curves of the first kind," *American Journal of Mathematics*, vol. 57 (1935), pp. 119-141.

$$\tilde{\boldsymbol{L}} = \boldsymbol{m}\tilde{\boldsymbol{Q}}, \tag{1}$$

where $\boldsymbol{m}$ is a numerical square matrix, in which (i) every element of the main diagonal is 1, (ii) every element to the left of the diagonal is 0, and (iii) certain elements to the right of the diagonal are -1, and the rest 0. The determinant $|\boldsymbol{m}| = 1$, so that we can also write

$$\tilde{\boldsymbol{Q}} = \boldsymbol{m}^{-1}\tilde{\boldsymbol{L}}. \tag{2}$$

We shall indicate the virtual intersection number of two curves or linear systems as a dot product; since all the curves Q_a have virtual grade -1, and no virtual intersections with each other, they have the intersection matrix

$$\tilde{\boldsymbol{Q}} \cdot \boldsymbol{Q} = -\boldsymbol{E},$$

where $\boldsymbol{E}$ indicates the unit or identical matrix; hence by (1) that of the curves L_a is

$$\tilde{\boldsymbol{L}} \cdot \boldsymbol{L} = \boldsymbol{m}\tilde{\boldsymbol{Q}} \cdot \boldsymbol{Q}\tilde{\boldsymbol{m}} = -\boldsymbol{m}\tilde{\boldsymbol{m}} = \boldsymbol{n} \tag{3}$$

say; the elements of the numerical matrices $\boldsymbol{m}$, $\boldsymbol{n}$, we shall write $m_{\alpha\beta}$, $n_{\alpha\beta}$.

If $m_{\alpha\beta} = -1$, the point O_β will be called proximate to O_α. (It is easy to see that this is equivalent to the definition of proximate points taken by Barber and Zariski from Enriques.) Since the curves L_a are irreducible, $n_{\alpha\beta}$ cannot be negative for $\alpha \neq \beta$, and we deduce at once the following results from (3):

If any point is proximate to two others, then one of these is proximate to the other; for if we had $m_{\alpha\gamma} = m_{\beta\gamma} = -1$, $m_{\alpha\beta} = 0$, $\alpha < \beta < \gamma$, then $n_{\alpha\beta}$ would be negative; similarly

No two points can be proximate to the same two others, since if we had $m_{\alpha\gamma} = m_{\beta\gamma} = m_{\alpha\delta} = m_{\beta\delta} = -1$, $\alpha < \beta < \gamma < \delta$, then even with $m_{\alpha\beta} = -1$, $n_{\alpha\beta}$ would be negative; finally

No point can be proximate to more than two others; for if we had $m_{\alpha\delta} = m_{\beta\delta} = m_{\gamma\delta} = -1$, $\alpha < \beta < \gamma < \delta$, then also $m_{\alpha\gamma} = m_{\beta\gamma} = m_{\alpha\beta} = -1$, and $n_{\alpha\beta}$ would be negative.

If $|K|$ be the canonical system on the surface, let $\boldsymbol{L} \cdot K = \boldsymbol{s}$, $\boldsymbol{Q} \cdot K = \boldsymbol{t}$; since the curves L_a, Q_a are all of genus 0, $t_a = -1$, and $s_a = -(2 + n_{aa})$. We must of course have

$$\tilde{\boldsymbol{s}} = \boldsymbol{m}\tilde{\boldsymbol{t}}; \tag{4}$$

but this also follows from (3), since from the form of $\boldsymbol{m}$ we have

$$-\sum_\beta (m_{\alpha\beta}) = \sum_\beta (m^2{}_{\alpha\beta}) - 2.$$

We now proceed to prove the general converse of these properties, or the generalisation of the well-known theorem:

(5) *If an irreducible curve has genus* 0 *and grade* -1, *then there exist regular systems, free from unassigned base points, on whose projective models the curve appears as the neighbourhood of a simple point.*

The theorem we shall prove is as follows:

(6) *If a set of irreducible curves all have genus* 0, *and if* (*for a suitable ordering of the curves*) *their intersection matrix can be put into the form* $-\boldsymbol{mm}$, *where* $\boldsymbol{m}$ *has the properties* (i), (ii), (iii), *above, then there exist regular systems, free from unassigned base points, on whose projective models the curves appear as the diminished neighbourhoods of a set of simple points.*

This we shall prove by induction, assuming it to be true for a set of $n-1$ curves, and hence proving it to be true for a set of n. The Theorem (5) itself provides the basis of induction, being precisely the case $n=1$.

We note first, since the last row of $\boldsymbol{m}$ consists entirely of 0's, except for a 1 in the last place, that the last column of $\boldsymbol{n}$ is the same as that of $\boldsymbol{m}$ with the sign changed; and in particular $n_{nn}=-1$; thus L_n is identical with Q_n, and is an exceptional curve as it stands. We can therefore, by (5), transform the surface Φ on which the curves are into a surface Φ^* on which L_n appears as the neighbourhood of a simple point O^*_n; the remaining curves L_a appear as a set of $n-1$ curves L^*_a, where L^*_a is the image of L_a or of L_a+L_n according as $n_{an}=0$ or 1. The intersection matrix of the curves $\boldsymbol{L^*}$ is $\boldsymbol{n^*}$, obtained from $\boldsymbol{n}$ by omitting the last row and column, and increasing by 1 each element $n_{a\beta}$ such that $n_{an}=n_{n\beta}=1$ (this just takes care of the increase of grade of those curves to which L_n is added, and of the possible intersection of two of the new curves in O^*_n). But on account of the identity (save for sign) of the last row and column of $\boldsymbol{n}$ with the last column of $\boldsymbol{m}$, this means that

$$\boldsymbol{n^*}=-\boldsymbol{m^*\tilde{m}^*},$$

where $\boldsymbol{m^*}$ is obtained from $\boldsymbol{m}$ by omitting the last row and column. Thus the curves L^*_a are a set to which the inductive hypothesis is applicable, and there exists a linear system, regular and free from unassigned base points on Φ^*, on whose projective model $\underline{\Phi}$ the curves L^*_a appear as the diminished neighbourhoods of a set of $n-1$ simple points O_a. This system is equally free from unassigned base points on the original surface Φ, and as it certainly

has not a base point at O^*_n the latter appears on $\underline{\Phi}$ as a simple point $\underline{O_n}$; the diminished neighbourhoods of the whole set of n points $\underline{O_a}$ are the images of the original set of curves L_a; thus the theorem is proved.

A square matrix is called reducible, if (possibly after a rearrangement of the rows, and corresponding rearrangement of the columns) it consists of a number of smaller square matrices arranged corner to corner down the main diagonal, and zeros everywhere else; these smaller matrices are called its constituents. If not reducible it is called irreducible. It is evident that $\boldsymbol{m}$, $\boldsymbol{n}$ are reducible or irreducible together.

We have the theorem:

If $\boldsymbol{m}$ and $\boldsymbol{n}$ are irreducible, the curves L_a are a simply connected tree (i. e., any two of them can be linked in an unique way by a chain of curves of the set, all different, and in which consecutive curves and only those intersect.) This also we prove by induction; for if $\boldsymbol{m}$ is irreducible, so is $\boldsymbol{m}^*$ and hence also $\boldsymbol{n}^*$; on the other hand if $\boldsymbol{n}$ is irreducible, L_n meets either one or two of the other curves, i. e., O^*_n is either a general point of one of the curves L^*_a, or the common point of two of them; and this being so, it is clear that if the curves L^*_a form a simply connected tree, so do the curves L_a.

Another thing that is evident if $\boldsymbol{m}$ and $\boldsymbol{n}$ are irreducible is that O_1 is the only one of the points O_a that is original, since if $n > 1$, O_n is certainly not original, being as we have seen proximate to at least one other point.

On the other hand, it is clear that if $\boldsymbol{m}$ be reducible, and $\boldsymbol{m}^{(i)}$ its constituents, those of $\boldsymbol{n}$ are $\boldsymbol{n}^{(i)} = -\boldsymbol{m}^{(i)}\tilde{\boldsymbol{m}}^{(i)}$, and that to each constituent of either corresponds a simply connected tree among the curves L_a, having no intersection with any of the other trees, and a subset of the points O_a consisting of one original point and other points in its various neighbourhoods; the different original points being at distinct points of Φ.

In conclusion, it may be pointed out that a general linear system $| C |$ on Φ corresponds to a system $| C |$ on $\underline{\Phi}$ having base points of multiplicity h_a at the points $\underline{O_a}$, where

$$\boldsymbol{h} = C \cdot \boldsymbol{Q} = C \cdot \boldsymbol{L}\tilde{\boldsymbol{m}}^{-1},$$

so that the linear system without base points on $\underline{\Phi}$ in which $| C |$ is totally contained is

$$\begin{aligned} | C + C \cdot \boldsymbol{Q}\tilde{\boldsymbol{Q}} | &= | C + C \cdot \boldsymbol{L}\tilde{\boldsymbol{m}}^{-1}\boldsymbol{m}^{-1}\tilde{\boldsymbol{L}} | \\ &= | C + C \cdot \boldsymbol{L}\boldsymbol{n}^{-1}\tilde{\boldsymbol{L}} |. \end{aligned}$$

Note. The criterion for the theorem (6), that the matrix $\boldsymbol{n}$ shall be expressible in the form $-\boldsymbol{m}\tilde{\boldsymbol{m}}$ is one that can be directly tested by constructing the matrix $\boldsymbol{m}$ if it exists. For since the last column of $\boldsymbol{m}$ is the

same as that of $\boldsymbol{n}$ with the sign changed, we can observe if $\boldsymbol{n}$ has any column consisting of -1 in the diagonal, and not more than two 1's elsewhere; if it has, we can rearrange the columns (and the rows correspondingly) so that this is the last, and take the last column of $\boldsymbol{m}$ to be the same with change of sign; we then construct the matrix $\boldsymbol{n}^*$ as before, and repeat the process to find the last column but one, and so on; remembering of course that every rearrangement of the rows and columns must be echoed as a rearrangement of the rows of the part of $\boldsymbol{m}$ already built up. If at any stage there is no diagonal element left which is -1, the process breaks down, and $\boldsymbol{m}$ does not exist; if we see a diagonal element -1 with more than two 1's in the same row or column, we can give up, as the process will certainly break down later on.

CAMBRIDGE, ENGLAND.

THEOREMS CONCERNING A BASIS FOR ALL SOLVABLE GROUPS, WITH A LIST OF BASIS GROUPS OF DEGREE < 64.

By A. C. Lunn and J. K. Senior.

In previous papers[1] it has been shown that *every solvable group of order* Πx_i $(x_i = p_i^{a_i},\ i = 1 \cdots n)$ *can be expressed and in one way only as a permutation group of degree* Σx_i *with* n *sets of transitivity of degrees respectively* x_i. Such expressions will here be called ΣU^* expressions. This result directs attention to a special category of permutation groups (to be called U^*-groups) having the following properties: Each U^* is transitive of some degree p^a and of some order Kp^a where K is prime to p, and is a solvable group containing no proper invariant subgroup of order dividing K. Clearly each transitive constituent of a ΣU^* expression is a U^*-group and conversely every U^*-group occurs as a transitive constituent of some ΣU^* expression, e. g. the ΣU^* expression of the abstract group (hereafter called a U-group) with which the U^*-group in question is simply isomorphic. It has been proven that if two U^*-groups are distinct as permutation groups, they cannot be simply isomorphic and hence there is a one to one correspondence between U^*- and U-groups. The category of U-groups is of interest since every solvable abstract group is either a U-group or can be formed from a properly chosen set of U-groups by direct multiplication or by the use of suitable isomorphisms. The U-groups thus constitute a basis, from the members of which all solvable groups may be constructed.

The one to one correspondence between U- and U^*-groups makes it convenient to identify U-groups by their U^*-expressions. A U^*-group of degree p^a and order Kp^a will be called $U^*(p^a, K)$. The purpose of the present paper is to consider certain properties of the U^*-groups and to give a complete list of the $U^*(p^a, K)$ groups where $p^a < 64$, $K > 1$ and $a > 1$. It will appear that where $K = 1$ or $a = 1$ the $U^*(p^a, K)$ groups belong to categories which are already too well known to require relisting.

Certain properties of U^*-groups which are useful in listing all those of given degree are shown in the two following theorems.[2] It has however been

[1] Lunn and Senior, *American Journal of Mathematics*, vol. 56 (1934), pp. 319, 511.

[2] In developing these theorems, we have been greatly assisted by valuable sug-

convenient to state these theorems in a form somewhat more general than that required for the immediate purpose. The following notation will be used.

(1) L and K are two relatively prime integers.

(2) If the order of a solvable group divides L or K, the group will be called respectively an L-group or a K-group.

(3) A solvable group whose order divides LK will be called a V^L-group if it contains no proper invariant K-subgroup; it will be called a V^K-group if it contains no proper invariant L-subgroup.

(4) $A(G)$ is defined as the group of automorphisms of G.
$B(G)$ " " " " " " inner " of G.
$O(G)$ " " " " " " quotient group $A(G)/B(G)$.

Note that

(a) Every L-group is a V^L-group and every K-group is a V^K-group, but not conversely.

(b) A solvable group G of order LK cannot be both a V^L-group and a V^K-group, since at least one of its sets of Sylow subgroups has an intersection invariant under G and of order > 1.

THEOREM I. *If G of order LK is a V^L-group:*

(a) *G uniquely determines each member of a series of characteristic subgroups of G such that $I = G_0 < G_1 \cdot \cdot \cdot < G_{n-1} < G_n = G$ and such that the quotient group G_{i+1}/G_i is the maximal invariant L-subgroup in G_n/G_i where i is even and the maximal invariant K-subgroup in G_n/G_i where i is odd.* This series will be called the V-series of G, and G (a V^L-group) will be written $V^L{}_n$, if it has $n-1$ proper subgroups in its V-series.

(b) *If G_i is a member of the V-series of G $(= V^L{}_n)$ then G_i is a $V^L{}_i$-group and the $i+1$ members of its V-series are identical respectively with the first $i+1$ members of the V-series of G.* Hence the V-series of G may be written

$$I = V^L{}_0 < V^L{}_1 \cdot \cdot \cdot < V^L{}_{n-1} < V^L{}_n = G.$$

A $V^L{}_{i+a}$ group will be said to be " stepped up " from a certain $V^L{}_i$-group when this $V^L{}_i$-group is the $(i+1)$-th member of the V-series of $V^L{}_{i+a}$.

Proof. All of the invariant L-subgroups of G generate the maximal

---

gestions from Mr. Philip Hall of Cambridge, England, and Mr. Garrett Birkhoff of Cambridge, Mass.

invariant L-subgroup of G. This characteristic subgroup of G is G_1. Since G is solvable, G/G_1 is solvable, and if $G > G_1$, then since G/G_1 contains no invariant L-subgroup, it contains at least one invariant K-subgroup. All the invariant K-subgroups of G/G_1 generate the maximal invariant K-subgroup of G/G_1. To this characteristic subgroup of G/G_1, there corresponds a characteristic subgroup of G which is G_2. By repetition of this argument, the V-series of G is completely determined.

Let F be any invariant subgroup of G. If F is an L-group, then it is also a V^L-group. If F is not an L-group and contains an invariant K-subgroup H, then H (being solvable) must contain at least one prime power subgroup J which is characteristic under H, hence invariant under F, and thus contained in the intersection D of a set of Sylow subgroups of F. This D would be a K-subgroup, characteristic under F and hence invariant under G, contrary to the definition of G as a V^L-group. Thus F is a V^L-group and in particular G_i is a V^L-group. (If F be a maximal invariant subgroup of G, successive repetitions of the above argument show that any subgroup which is a member of any ordinary series of decomposition of G is a V^L-group.)

In G_i, let the maximal invariant L-subgroup be G_{i1}. G_{i1} is characteristic under G_i which is characteristic under G. Hence G_{i1} is characteristic under G. G_{i1} contains G_1, but as G_1 is the maximal invariant L-subgroup of G, $G_{i1} = G_1$. Thus the first two members of the V-series of G_i are identical respectively with the first two members of the V-series of G. Let Q_{i2} be the largest invariant K-subgroup in G_i/G_1. To this characteristic subgroup of G_i/G_1 there corresponds G_{i2}, a characteristic subgroup of G_i which is also characteristic under G. G_{i2} contains G_2, but since G_2 is the largest invariant subgroup of G containing G_1 in such wise that G_i/G_1 is a K-group, then $G_{i2} = G_2$ and the third member of the V-series of G_i is identical with the third member of the V-series of G. Repetition of the above argument proves that where $j \gtreqless i+1$, the j-th member of the V-series of G_i is identical with the j-th member of the V-series of $G_n = G$. The proof of Theorem I is thus complete.

THEOREM II. *If G of order LK is a $V^L{}_n$-group, then*

(a) *When m is even*

(1) *To $V^L{}_{m+1}/V^L{}_{m+1}$ corresponds a set of conjugate V^K-subgroups in $O(V^L{}_{m+1}/V^L{}_m)$*
(2) *To $V^L{}_{m+2}/V^L{}_{m+1}$ " " " " " K-subgroups in $A(V^L{}_{m+1}/V^L{}_m)$*

(b) *When m is odd*

(1) *To $V^L{}_{m+1}/V^L{}_{m+1}$ corresponds a set of conjugate V^L-subgroups in $O(V^L{}_{m+1}/V^L{}_m)$*
(2) *To $V^L{}_{m+2}/V^L{}_{m+1}$ " " " " " L-subgroups in $A(V^L{}_{m+1}/V^L{}_m)$.*

Here the term "corresponds" is used in a familiar special sense to imply not only abstract isomorphism but an appropriate equivalence of automorphism; the exact nature of this relation will appear in the course of the proof.

Proof. If m is even, $V^{L_{m+1}}$ is the maximal invariant subgroup of G such that $V^{L_{m+1}}/V^{L_m}$ is an L-group. Hence $V^{L_{m+2}}/V^{L_{m+1}}$ is a K-group and $V^{L_{m+4}}/V^{L_{m+1}}$ contains no invariant L-subgroup and is thus a V^K-group. Similarly when m is odd $V^{L_{m+2}}/V^{L_{m+1}}$ is an L-group and $V^{L_{m+4}}/V^{L_{m+1}}$ is a V^L-group. To complete the proof of Theorem II it will now be shown that

When $m = 0$, *and* $V^{L_m} = V^{L_0} =$ *identity and* $i > 1$, *then*

V^{L_i}/V^{L_1} *corresponds to a set of conjugate* V^K*-subgroups in* $O(V^{L_1})$ *and*
V^{L_2}/V^{L_1} " " " " " " *K-subgroups in* $A(V^{L_1})$.

Let J be the subgroup of V^{L_i} which consists of all the operators in V^{L_i} which are commutative with every operator in V^{L_1}. Since V^{L_1} is a characteristic subgroup of V^{L_i}, both the central C of V^{L_1} and the "centralizer" J of V^{L_1} in V^{L_i} are characteristic under V^{L_i}. The intersection of V^{L_i} and J is C. Consider the quotient group J/C. Since V^{L_i} is solvable, J/C is solvable, and the intersection D of some set of Sylow subgroups in J/C is characteristic under J/C. Obviously D is either an L-group or a K-group, if the order of J/C is > 1. Let E be the invariant subgroup of V^{L_i} which corresponds to D in J/C. Suppose D to be an L-group. Then E would be an L-group since C (contained in V^{L_1}) is an L-group, and E would generate with V^{L_1} an invariant L-subgroup of V^{L_i} whose order would exceed that of V^{L_1}, contradicting the definition of V^{L_1}; hence D is not an L-group. Suppose D to be a K-group. Since every operator in E (a subgroup of J) is commutative with every operator in C (a subgroup of V^{L_1}), then E would be the direct product of C and F, the Sylow K-subgroup of E. Since E would be invariant under V^{L_i}, and F would be characteristic under E, F would be an invariant K-subgroup of V^{L_i}, contrary to the definition of V^{L_i}; hence D is not a K-group. Since D is neither a K-group nor an L-group, J/C is of order 1 and $J = C$. It follows that every operator of V^{L_i} not in C sets up an automorphism in V^{L_1}.

Since V^{L_i}/C is the group of all automorphisms set up in V^{L_1} by V^{L_i}, then V^{L_i}/C "corresponds" to a set of conjugate subgroups in $A(V^{L_1})$. V^{L_1}/C corresponds to $B(V^{L_1})$ and hence V^{L_i}/V^{L_1} corresponds to a set of conjugate subgroups in $A(V^{L_1})/B(V^{L_1}) = O(V^{L_1})$. Since V^{L_i}/V^{L_1} is solvable and contains no invariant L-subgroups, the corresponding set of conjugate subgroups in $O(V^{L_1})$ must be V^K-groups.

Since V^{L_2} is solvable and the order of V^{L_1} is prime to the index of V^{L_1} under V^{L_2}, then V^{L_2} contains exactly one set of conjugate K-subgroups simply isomorphic with V^{L_2}/V^{L_1}.[3] As no operator in such a K-subgroup is contained in C, every such operator sets up an automorphism in V^{L_1}. Hence such a K-subgroup is a subgroup of $A(V^{L_1})$ and V^{L_2}/V^{L_1} corresponds to this K-subgroup or to one of its conjugates under $A(V^{L_1})$. This completes the proof of Theorem II.

The aspect of Theorems I and II which is of particular use in the listing of U^*-groups may be briefly stated as follows:

In any V series V_{i+2}/V_{i+1} is of order relatively prime to the order of V_{i+1}/V_i and corresponds to a set of conjugate subgroups in $A(V_{i+1}/V_i)$.

If (V_{i+1}/V_i) is cyclic and of odd order, $G = V_n$ is either V_{i+1} or V_{i+2}. If $A(V_{i+1}/V_i)$ contains no operator of order prime to the order of (V_{i+1}/V_i) then $G = V_n = V_{i+1}$. This is the case (for instance) where V_{i+1}/V_i is cyclic and of order 2^v. Theorem I indicates a method for determining all the V^L-groups stepped up from a given V^{L_1}. First determine all the V^{L_2}-groups stepped up from the given V^{L_1}; then determine all the V^{L_3}-groups stepped up from these V^{L_2}-groups; etc. That the entire list of V^L-groups thus obtained from a single V^{L_1} comprises only a finite number of groups is a consequence of Theorem II.

The application of Theorems I and II where L is a power of a prime. If G is a solvable V^L-group of order LK it contains (according to Hall's theorem) a single set of conjugate subgroups of order K and can thus be expressed and in one way only as a transitive group of degree L. Where L is a power of a prime, G is thus a U-group as defined at the beginning of this paper, and its transitive expression of degree L is a U^*-group. In order to emphasize certain special properties of these groups it is convenient to handle them with a distinctive notation. Accordingly, when L is a power of a prime, V will be replaced by U. The symbol V^{L_i} becomes $U_i(p^a)$ and the U-series of $U_n(p^a)$ may be written,

$$1 = U_0(p^a) < U_1(p^a) \cdots < U_{n-1}(p^a) < U_n(p^a) = G.$$

In the sense of this notation, a solvable group G is a U-group if some one prime factor of g divides the order of every proper invariant subgroup. As previously shown, no solvable group can have this relation to more than one

[3] P. Hall, *Journal of the London Mathematical Society*, vol. 3 (1932), p. 98.

prime. The equation $G = U_n(p^a)$ means that G is a U-group with respect to the prime p and that it has $n-1$ proper subgroups in its U-series. If the fuller notation $U_n(p^a, K)$ is used, it will be understood in addition that G is precisely of order p^aK with K prime to p.

According to this notation if $K = 1$, U is a prime power group U_1, and U^* is its regular expression. Where $a = 1$, U^* is either the metacyclic group of degree p and order $p(p-1)$ or one of its invariant subgroups, since no other solvable transitive groups of prime degree exist. Both of these categories, the prime power groups and the metacyclic groups with their invariant subgroups are too well known to require discussion here.

In $U_n(p^a, K)$ the U-series determines a sequence of integers $a_1 \cdots a_f$ such that $\sum_1^f a_i = a$ and p^{a_i} is the index of U_{2i-2} under U_{2i-1}. If n is even, $f = n/2$; if n is odd $f = (n+1)/2$. This sequence of integers will be called the U-sequence of U. Hall's results [4] show that if F is a group of order p^y, the order of the p-Sylow subgroup in $A(F)$, and *a fortiori* the order of the p-Sylow subgroup in $O(F)$ does not exceed $p^{y(y-1)/2}$. But U_{2i-1}/U_{2i-2} is a p-group and U_n/U_{2i-1} corresponds to a set of conjugate subgroups in $O(U_{2i-1}/U_{2i-2})$. Hence the U-sequence of U_n is subject to the restriction that $\sum_{i+1}^f a_j \leqq a_i(a_i-1)/2$. It follows that the sequence $a_1 \cdots a_f$ dominates ($\geqq$) the arithmetic sequence $f \cdots 2, 1$. Hence if the number t is in the U-sequence, it is one of the last t members. However, it seems improbable that in the case of every sequence satisfying the above restriction, there exists a group for which it is the U-sequence.

The problem of constructing all the groups "stepped up" from a given U_1 may be attacked in various ways. The methods outlined below have proven convenient in practice and adequate for U^*-groups of degrees < 64. They are based primarily on the consideration that

(a) *Where i is even (U_{i+2}/U_{i+1}) corresponds to a set of conjugate K-subgroups in $A(U_{i+1}/U_i)$.*

(b) *Where i is odd (U_{i+2}/U_{i+1}) corresponds to a set of conjugate p^w-subgroups in $A(U_{i+1}/U_i)$.*

The construction of all possible U_2 groups stepped up from a given U_1 is easy since there is exactly one such group for every set of conjugate K-sub-

[4] P. Hall, *Proceedings of the London Mathematical Society*, ser. 2, vol. 36 (1933-34), pp. 36, 37.

groups in $A(U_1)$. U^*_2 is thus always a subgroup of the holomorph of U^*_1. The search for all the U_2 groups stepped up from all possible U_1 groups of given order is much facilitated by a theorem of Hall. His results [5] show that if F is a prime power group and E the elementary group of the same order as F, then every K-subgroup in $A(F)$ is simply isomorphic with a K-subgroup in $A(E)$.

The problem of constructing all possible U_3 groups stepped up from a given U_2 is somewhat more intricate. Let M_2 be a maximal K-subgroup of U_2. M_2 is simply isomorphic to U_2/U_1, and with U_1 generates U_2. Each M_2 is also a maximal K-subgroup of U_3 and contained in its invariant subgroup U_2. If N_2 and N_3 are the normalizers of M_2 in U_2 and U_3 respectively then N_3/N_2 is simply isomorphic with U_3/U_2. Since U_3/U_2 corresponds to a set of conjugate p-subgroups in $A(U_2/U_1)$ [and therefore in $A(M_2)$], U_3 may be generated by adjoining to U_2 generators $j_1 \cdots j_s$ which

(a) generate with U_1 a p-Sylow subgroup of U_3,

(b) transform the operators of M_2 according to the group (U_3/U_2).

If h_i is the lowest power of j_i contained in U_1, then h_i must set up the identity automorphism in M_2. Thus h_i is an operator in H_2, the subgroup of U_1 which consists of all those operators in U_1 which are commutative with every operator in M_2. Let C_i be the central of U_i. Since every operator in U_i which is not in C_1 sets up an automorphism in U_1, $C_1 \leqq C_i$. Thus H_2 always contains C_2 and it is convenient to distinguish two categories of cases (1) $H_2 = C_2$, (2) $H_2 > C_2$.

(1) $H_2 = C_2$. This case always arises when U_1 is abelian, since under these circumstances, any operator in U_1 commutative with every operator in M_2 is in C_2. But there are numerous cases where $H_2 = C_2$ although U_1 is not abelian. In all cases of this category however, h_i sets up the identity automorphism in U_1 and U_3/U_1 corresponds to a set of conjugate V^K-subgroups in $A(U_1)$. It is convenient to divide this category of U_2 groups according to whether (a) $H_2 = C_2 = 1$ or (b) $H_2 = C_2 > 1$.

(a) $H_2 = C_2 = 1$. Here M_2 and $j_1 \cdots j_s$ generate a V^K-subgroup of $A(U_1)$ simply isomorphic with U_3/U_1 and there is a one to one correspondence between the U_3 groups derived from a given U_2 and the sets of conjugate V^K-subgroups in $A(U_1)$ containing the given M_2 (or one of its conjugates) as an invariant subgroup of index p^v. Every such U_3 group may be expressed not only as a U^*-group but as a subgroup of the holomorph of its U^*_1.

---

[5] P. Hall, *Proceedings of the London Mathematical Society*, ser. 2, vol. 36 (1933-34), pp. 37, 38.

(b) $H_2 = C_2 > 1$. In this case, a variety of abstract groups U_3 is sometimes obtainable from the same U_2 and the same subgroup of automorphisms of U_1. To choose the j_i's so that each $h_i = 1$ is always possible; if they are so chosen, the resulting U_3 may be written as a subgroup of the holomorph of U^*_1 exactly as in the case where $H_2 = C_2 = 1$. The generators so chosen will be called j'_i's, and the U_3-group thus generated will be called a principal U_3-group. But where $H_2 = C_2 > 1$, it is frequently possible to choose the j_i's so that they set up respectively in U_2 the same automorphisms as the j'_i's but so that not every h_i is identity. With generators so chosen the resulting U_3 is sometimes (not always) distinct from the principal U_3 group which corresponds to the same set of conjugate V^K-subgroups in $A(U_1)$. Such a distinct U_3-group will be called a collateral U_3-group; it cannot be written as a subgroup of the holomorph of its U^*_1.

It is thus shown that when $H_2 = C_2 > 1$, the correspondence between U_3-groups and the suitable sets of V^K-subgroups in $A(U_1)$ is not always one to one but in certain cases many to one. In an extended program for the construction of U_3 groups it would be desirable to develop rules for the number of collateral U_3 groups belonging to any principal U_3. But the subject is an intricate one and, within the field covered by this paper, the cases where collateral groups arise occasion no great difficulty and have been worked out individually.

(2) $H_2 > C_2$. In this case U_3/U_1 may correspond to a set of conjugate V^K-subgroups not in $A(U_1)$ but in $O(U_1)$. In a general program, the U_3 groups stepped up from U_2 groups of this category would require detailed treatment, but inspection of all U_2 groups of low order has shown that the relation $H_2 > C_2$ is impossible if U_2 is of order < 96. A U_3 group stepped up from this sort of U_2 would have an order $\leqq 192$ and a degree $\leqq 64$. Further discussion of this case is therefore omitted.

Considerations similar to those already given hold for the construction of U_n groups where $n > 3$. But sufficient analysis has been carried out to show that, for degrees < 64, there is only one instance of a U^*_n where $n > 3$; so that no further discussion of the general theory here required is now attempted.

The U^* groups of degrees < 64 and their corresponding U groups. If U^* is of order p^aK and degree $p^a < 64$ *then* $a < 6$ and the presumptive U-sequences are as follows:

a	*U-sequences*
1	1
2	2
3	3 2, 1
4	4 3, 1
5	5 4, 1 3, 2.

But consider the sequence 3, 2. Here p must be 2 and U_1 of order 8. $O(U_1)$ must contain a set of conjugate solvable subgroups of order $4K$ (where $K > 1$) which contain no invariant subgroup of order 2 or 4. If G is of order 8, no $O(G)$ meets this condition and so, where U^* is of degree < 64, the U-sequence 3, 2 does not occur.

If the cases where $a = 1$ or $K = 1$ are omitted (for reasons already explained), and if it is borne in mind that, when U_{i+1}/U_i is of order 2, then $U_{i+1} = U_n$, the following table gives the possibilities for a U^*-group of degree < 64.

a	*U-sequence*	p	n *in* U_n
2	2	2 or 3 or 5 or 7	2
3	3	2 or 3	2
	2, 1	2	3
		3	3 or 4
4	4	2	2
	3, 1	2	3
5	5	2	2
	4, 1	2	3.

From the table it appears that, where $p > 2$, n exceeds 2 only where U^* is of degree 27 and U_1 of order 9 and therefore abelian. U_3/U_1 and U_4/U_1 in this case must correspond to subgroups of $A(U_1)$ which are of order $3K$ but contain no invariant subgroup of order 3. Hence U_1 is not cyclic. If U_1 is not cyclic and of order 9, the only subgroups of $A(U_1)$ which have orders divisible by 3 but contain no invariant subgroups of order 3 are $A(U_1)$ itself and its

single subgroup of index 2. If U_n/U_1 corresponds to either of these groups, U_2 is such that $H_2 = C_2 = 1$ and the correspondence between the U_n-groups and the subgroups of $A(U_1)$ is one to one. Thus, if $p > 2$ and $n > 2$, there are only two U^*-groups of degree < 64. Of the corresponding U-groups one $(n = 4)$ may be written as the holomorph (of order 432) of the non-cyclic group of order 9; the other $(n = 3)$ may be written as the positive half of this holomorph.

Where $p = 2$, n never exceeds 3 and attains this figure only for the sequences 2, 1 or 3, 1 or 4, 1.

(a) If U_1 is of order 4, $K = 3$, and U_3 is of order 24 and a principal U_3 group.

(b) If U_1 is of order 8, $K = 3$, or 7 or 21 and U_3 would be of order 48, 112 or 336. But in the latter two cases U_3/U_1 would correspond to a set of conjugate subgroups of order 14 or 42 in $A(U_1)$ where U_1 is the elementary group of order 8. As this $A(U_1)$ contains no subgroups of orders 14 or 42, U_3 is of order 48, where U_1 is of order 8. Since, in the U_2 groups of order 24, $H_2 = C_2 > 1$, the possibility of collateral groups of order 48 must be considered.

(c) If U_1 is of order 16, $K = 3, 5, 7, 9, 15$, or 21. If $K > 3$, U_1 is the elementary group of order 16 and U_3/U_1 corresponds to a set of conjugate subgroups of order $2K$ in $A(U_1)$. As this $A(U_1)$ contains no subgroups of order 14 or 42 there are no U_3 groups of order 224 or 672. Where $K = 5$, 9 or 15, U_2 is such that $H_2 = C_2 = 1$ and no collateral U_3 groups of orders 160, 288 or 480 are possible. Where $K = 3$, U_1 is not necessarily elementary and U_2 is sometimes such that $H_2 = C_2 > 1$, so that collateral groups of order 96 are to be looked for.

It is thus shown that if U^*_n is of degree < 64, $n > 2$ only for the orders 24, 48, 96, 160, 216, 288, 432 and 480, and that, for orders other than 48 and 96, collateral U_3 groups are impossible. A U_4 group occurs only in the order 432. As all the groups of orders 48 and 96 are known, a discussion of the number of collateral groups is superfluous and the list of groups of these orders is taken from the published tables. In all other orders the correspondence between U_n groups $(n > 1)$ and the suitable sets of conjugate subgroups in $A(U_1)$ is one to one.

The following table summarizes the U^*-groups of degree < 64 where $a > 1$ and $K > 1$

		Orders of		
Degree	*K*	U_2	U_3	U_4
4	3	12	..	...
8	Factors of 21	$8K$	24	...
9	Factors of 16	$9K$	..	...
16	3, 5, 7, 9, 15, 21	$16K$	48	...
25	Factors of 96	$25K$	..	...
27	2, 4, 8, 13, 16, 26, 32	$27K$	216	432
32	3, 5, 7, 9, 15, 21, 31, 63, 155	$32K$	96, 160	
			288, 480	...
49	Proper factors of 288	$49K$	..	...

As appears from the following list there are just 212 of these groups. To these are added the 101 U^*-groups of degree < 64 where $a = 1$, and the 83 U^*_1-groups of degree < 64 where $a > 1$. Thus the basis groups whose U^* expressions are of degree < 64 are 396 in number. Any solvable group which contains no Sylow subgroup of order $\geqq 64$ may be constructed from these 396 groups. The number of such groups is hard to estimate, but, if the basis is limited to the 102 U^*_1-groups in the above list, the calculation is easy and shows that approximately 1.75×10^8 groups can be constructed from this portion of the basis groups. There can therefore be little doubt that the number of groups which can be constructed from all the 396 basis groups of degree < 64 easily exceeds 10^{10}.

The Non-regular U^*-groups of degrees 4, 8, 9, 16, 25, 27, 32 *and* 49. The authors have proven the completeness of the following list—e. g. they have shown that where $p = 5$ and G is the non-cyclic group of order p^2, $A(G)$ contains exactly 29 sets of conjugate solvable K-subgroups. But these arguments are too lengthy for inclusion here.

The generators of the various U^*-groups are given in a particular order. The symbol X_x stands for a generator, and

$$A_1 \cdot \cdot \cdot A_a \mid B_1 \cdot \cdot \cdot B_b \mid C_1 \cdot \cdot \cdot M_m \mid N_1 \cdot \cdot \cdot N_n$$

means that the generators $A_1 \cdot \cdot \cdot A_a$ generate U_1; the generators $B_1 \cdot \cdot \cdot B_b$ extend U_1 to U_2; $C_1 \cdot \cdot \cdot C_c$ extend U_2 to U_3; etc. Furthermore the generators are arranged in such order that the group generated by the first x generators is invariant under the group generated by the first $x + 1$ generators.

In the degrees 16 and 32, each principal U_3 group of order 48 or 96 respectively, if it has collateral groups, is labelled '*pr*' and bracketed with its collateral groups labelled '*co.*'

DEGREE 4

	U*				ORDER
A_1 = ac.bd	#1	A_1A_2	B_1		12
A_2 = ab.cd					
B_1 = cbd					

DEGREE 8

	U*				ORDER
A_1 = abcd.efgh	#1	A_1A_2	B_1		24
A_2 = aecg.bhdf	#2	$A_3A_4A_5$	B_3		24
A_3 = ac.bd.eg.fh	#3	·	B_4		56
A_4 = ab.cd.ef.gh	#4	·	B_2B_3		168
A_5 = ae.bf.cg.dh	#5	A_1A_4	B_3	C_1	24
B_1 = bef.dgh					
B_2 = ecbgdhf					
B_3 = cbd.gfh					
C_1 = ae.bh.cg.df					

DEGREE 9

	U*			ORDER
A_1 = adh.beicfg	#1	A_1	B_1	18
A_2 = abc.def.ghi	#2	A_2A_3	B_2	18
A_3 = adg.beh.cfi	#3	·	B_3	18
B_1 = bc.dg.ei.fh	#4	·	B_1B_2	36
B_2 = dg.eh.fi	#5	·	B_3	36
B_3 = bdcg.efih	#6	·	B_3B_2	72
B_4 = beci.dhgf	#7	·	B_1B_4	72
B_5 = bhdecfgi	#8	·	B_5	72
	#9	·	B_3B_2	44

DEGREE 16

	U*				ORDER
A_1 = ac.bd.eg.fh.ik.jl.mo.np	#1	$A_1A_2A_4$	B_2		48
A_2 = ab.cd.ef.gh.ij.kl.mn.op	#2	$A_3A_1A_4$	B_4		48
A_3 = ae.bf.cg.dh.im.jn.ko.lp	#3	$A_3A_1A_2$	B_4		48
A_4 = ai.bj.ck.dl.em.fn.go.hp	#4	A_1A_4	B_3		48
A_5 = abcd.efgh.ijkl.mnop	#5	$A_1A_2A_3A_4$	B_2		48
A_6 = aeim.bfjn.cgko.dhlp	#6	·	B_3		48
A_7 = aecg.bhdf.imko.jpln	#7	·	B_4		80
A_8 = aick.bjdl.emgo.fnhp	#8	·	B_1		112
B_1 = ecbgdhf.mkjolpn	#9	·	B_1B_2		144
B_2 = cbd.gfh.kjl.onp	#10	·	B_4B_1		240
B_3 = cbd.eim.flo.gjp.hkn	#11	·	B_1B_2		336
B_4 = bef.dgh.jmn.lop	#12	$A_1A_2A_3$	B_2	C_1 pr	48
B_5 = bep.cik.dmf.ghl.jon	#13	·	·	C_2 co	48
B_6 = cplkm.bgone.djfhi	#14	A_5A_7	B_4	C_3 pr	48
C_1 = ai.bl.ck.dj.emfp.gohn	#15	·	·	C_4 co	48
C_2 = aiem.blfp.ckgo.djhn					
C_3 = ai.bl.ck.dj.epfo.gnhm					
C_4 = aick.bldj.epgn.fohm					

DEGREE 25

	U*			ORDER	U*			ORDER
A_1=aforvbgkswchltxdimpyejnqu	#1	A_1	B_1	50	#17	A_2A_3	B_{10}	200
A_2=abcde.fghij.klmno.pqrst.uvwxy	#2	"	B_5	100	#18	"	$B_3B_1B_2$	300
A_3=afkpu.bglqv.chmrw.dinsx.ejoty	#3	A_2A_3	B_1	50	#19	"	B_3B_5	300
B_1=be.cd.fu.gy.hx.iw.jv.kp.lt.ms.nr.oq	#4	"	B_2	50	#20	"	B_3B_4	300
B_2=fu.gv.hw.ix.jy.kp.lq.mr.ns.ot	#5	"	B_3	75	#21	"	$B_{10}B_2$	400
B_3=bwh.cto.dlq.eix.fmn.gjp.kyv.rus	#6	"	B_1B_2	100	#22	"	$B_4B_7B_5$	400
B_4=bxg.cqm.ehy.dos.fir.ptv.uwn.klj	#7	"	B_5	100	#23	"	B_5B_8	400
B_5=bced.fkup.gmys.hoxq.ilwt.jnvr	#8	"	B_6	100	#24	"	B_3B_{10}	600
B_6=bced.fpuk.xlht.qwoi.gryn.jsvm	#9	"	B_8	100	#25	"	B_3B_{11}	600
B_7=bfeu.ckdp.xqho.litw.gjyv.rmns	#10	"	B_9	100	#26	"	$B_3B_5B_2$	600
B_8=bced.ghji.lmon.qrts.vwyx	#11	"	B_3B_1	150	#27	"	$B_6B_7B_4$	600
B_9=fu.kp.bced.gwjx.hyiv.lros.mtnq	#12	"	B_3B_2	150	#28	"	$B_{10}B_2B_8$	800
B_{10}=bpcfekdu.hjonxvqr.iylswgtm	#13	"	B_6B_1	200	#29	"	$B_3B_{10}B_2$	1200
B_{11}=bkdfepcu.hyqmxgos.ijtrwvlr	#14	"	B_5B_2	200	#30	"	$B_6B_7B_4B_5$	1200
	#15	"	B_7B_2	200	#31	"	$B_6B_7B_4B_3D_{10}$	2400
	#16	"	B_6B_7	200				

DEGREE 27

	U*			ORDER	U*			ORDER
A_1=ajvdmzhqtbkwenαiruclxfoygps	#1	A_1	B_1	54	#23	$A_3A_4A_5$	B_{13}	216
A_2=adhbeicfg.jmqknrlop.svztwαuxy	#2	A_2A_5	B_1	54	#24	"	B_{14}	216
A_3=abc.def.ghi.jkl.mno.pqr.stu.vwx.yzα	#3	"	B_2	54	#25	"	B_7B_3	216
A_4=adg.beh.cfi.jmp.knq.lor.svy.twz.uxα	#4	"	B_3	54	#26	"	B_7B_5	216
A_5=ajs.bkt.clu.dmv.enw.fbx.gpy.hqz.irα	#5	"	B_2B_3	108	#27	"	B_7B_4	216
A_6=ajs.bkt.clu.dnx.eov.fmw.grz.hpα.iqy	#6	A_2A_6	B_2	54	#28	"	B_4B_5	216
B_1=bc.dg.ei.fh.js.ku.lt.my.na.oz.pv.qx.rw	#7	$A_3A_4A_6$	B_2	54	#29	"	B_7B_8	216
B_2=bc.dg.ei.fh.kl.mp.nr.oq.tu.vy.wa.xz	#8	"	B_4	54	#30	"	B_7B_{10}	216
B_3=js.kt.lu.mv.nw.ox.py.qz.rα	#9	"	B_2B_4	108	#31	"	$B_2B_3B_5$	216
B_4=dg.eh.fi.js.kt.lu.my.nz.oα.pv.qw.rx	#10	"	B_{11}	108	#32	"	B_{16}	351
B_5=dg.eh.fi.mp.nq.or.vy.wz.xα	#11	"	B_{15}	216	#33	"	$B_{13}B_3$	432
B_6=bc.ef.hi.js.ku.lt.mv.nx.ow.py.qα.rz	#12	"	$B_{11}B_6$	216	#34	"	$B_7B_5B_3$	432
B_7=bdcg.efih.kmlp.norq.tvuy.wxαz	#13	"	$B_{11}B_{12}$	216	#35	"	$B_7B_8B_3$	432
B_8=beci.dhgf.knlr.mqpo.twuα.vzyx	#14	"	$B_{15}B_6$	432	#36	"	$B_{13}B_5$	432
B_9=js.bdcg.efih.kvly.mupt.nxrz.oα.qw	#15	$A_3A_4A_5$	B_1	54	#37	"	$B_{13}B_9$	432
B_{10}=js.beci.dhgf.kwlα.mzpx.tnur.vqyo	#16	"	B_2	54	#38	"	$B_{14}B_5$	432
B_{11}=djgs.oqαw.ekht.mryx.fliu.npzv	#17	"	B_3	54	#39	"	$D_{16}B_4$	432
B_{12}=dogα.jwsq.emhy.kxtr.fniz.lvup	#18	"	B_2B_3	108	#40	"	$B_{16}D_1$	702
B_{13}=beghcidf.kupqlrmo.twyzuαvx	#19	"	B_2B_4	108	#41	"	$B_{13}B_5B_3$	864
B_{14}=js.beghcidf.kwpzlamx.tnyqurvo	#20	"	B_2B_5	108	#42	A_3A_4	$B_7B_{13}C_1$	216
B_{15}=bc.dwjogqsα.evknhptz.fxlmiruy	#21	"	B_7	108	#43	A_3A_4	$B_7B_{13}C_1D_1$	432
B_{16}=bdjemnqzufpwl.cgsiyαxokhvrt	#22	"	B_4	108				
C_1=ajs.bkt.clu.dnx.eov.fmw.grz.hpα.iqy								
D_1=dg.eh.fi.js.kt.lu.my.nz.oα.pv.qw.rx								

DEGREE 32

	U*				ORDER
A_1=ac.bd.eg.fh.ik.jl.mo.np.AC.BD.EG.FH.IK.JL.MO.NP	#1	$A_1\,A_2\,A_7$	B_1		96
A_2=ab.cd.ef.gh.ij.kl.mn.op.AB.CD.EF.GH.IJ.KL.MN.OP	#2	$A_1\,A_2\,A_6\,A_5$	B_1		96
A_3=ae.bf.cg.dh.im.jn.ko.lp.AE.BF.CG.DH.IM.JN.KO.LP	#3	.	B_2		224
A_4=ai.bj.ck.dl.em.fn.go.hp.AI.BJ.CK.DL.EM.FN.GO.HP	#4	.	$B_2\,B_1$		672
A_5=aA.bB.cC.dD.eE.fF.gG.hH.iI.jJ.kK.lL.mM.nN.oO.pP	#5	$A_1\,A_2\,A_6\,A_8$	B_1		96
A_6=aeim.bfjn.cgko.dhlp.AEIM.BFJN.CGKO.DHLP	#6	$A_1\,A_2\,A_3\,A_4\,A_5$	B_1		96
A_7=aAeEiImM.bBfFjJnN.cCgGkKoO.dDhHlLpP	#7	.	B_3		96
A_8=aAbBcCdDeMfNgOhP.iIjJkKlLmEnFoGpH	#8	.	B_4		160
A_9=abcdefghijklmnop.ABCDEFGH.IJKLMNOP	#9	.	B_2		224
A_{10}=aecg.bhdf.imko.jpln.AECG.BHDF.IMKO.JPLN	#10	.	$B_1\,B_3$		288
A_{11}=aickbjdl.emgo.fnhp.AICK.BJDL.EMGO.FNHP	#11	"	$B_3\,B_4$		480
A_{12}=aAiIcCkK.bBjJdDlL.eEmMgGoO.fFnNhHpP	#12	"	$B_2\,B_1$		672
A_{13}=aAiI.bBjJ.cCkK.dDlL.eEmM.fFnN.gGoO.hHpP	#13	.	$B_2\,B_5$		672
A_{14}=aAcCbBdD.eEgG.fFhHiIkK.jJlL.mMoO.nNpP	#14	"	$B_2\,B_3$		672
A_{15}=aA.bB.cC.dD.eE.fF.gG.hH.iK.jL.kI.lJ.mO.nP.oM.pN	#15	.	B_6		992
A_{16}=aA.bD.cC.dB.eM.fP.gO.hN.iI.jL.kK.lJ.mE.nH.oG.pF	#16	.	$B_2\,B_1\,B_5$		2016
A_{17}=aAbJcCdLeQfMgHhFiIjBkKlDmGnPoEpN	#17	.	$B_6\,B_4$		4960
A_{18}=aAbIcEdMeCfKgGhOiBjJkFlNmQnLoHpP	#18	$A_9\,A_{10}\,A_{12}$	B_7		96
A_{19}=aAiI.bFjN.cCkK.dHlP.eEmM.fBnJ.gGoO.hDpL	#19	$A_9\,A_{10}\,A_{11}\,A_5$	B_7		96
B_1=cbd.gfh.kjl.onp.CBD.GFH.KJL.ONP	#20	$A_9\,A_{10}\,A_{13}$	B_7		96
B_2=AcbCdDB.EgfGhHF.IkjKlLJ.MonOpPN	#21	$A_9\,A_{10}\,A_4\,A_3$	B_7		96
B_3=cbd.ei m.gj p.hkn.flo.CBD.EIM.GJP.HKN.FLO	#22	.	B_8		96
B_4=cpl km.bgo ne.djf hi.CPL KM.BGO NE.DJF HI	#23	.	B_9		96
B_5=eim.fjn.gko.hlp.EIM.FJN.GKO.HLP	#24	.	$B_7\,B_9$		288
B_6=AbgLoGKinEOHMDdcePBhNFImCfJkjlp	#25	$A_9\,A_{10}\,A_{11}\,A_{14}$	B_7		96
B_7=bef.dgh.jmn.lop.BEF.DGH.JMN.LOP	#26	.	B_9		96
B_8=iAI.jBJ.kCK.lDL.mEM.nFN.oGO.pHP	#27	.	$B_7\,B_9$		288
B_9=bef.dgh.iAI.jEN.kCK.lGP.mFJ.nBM.oHL.pDO	#28	$A_9\,A_{10}\,A_4\,A_{15}$	B_7		96
B_{10}=jmnAK.lopCI.bPkBf.dNiDh.eMFL.GgOHJE	#29	.	B_{10}		160
B_{11}=bep.cik.dmf.ghl.jon.BEP.CIK.DMF.GHL.JON	#30	$A_9\,A_4\,A_5$	B_{11}		96
B_{12}=bAP.eio.dIB.fEL.gmk.hMF.jON.lGD.nKJ.pCH	#31	$A_9\,A_4\,A_{16}$	B_{11}		96
C_1=aA.bD.cC.dB.eE.fH.gG.hF.iI.jL.kK.lJ.mM.nP.oO.pN	#32	$A_9\,A_4\,A_{17}$	B_{11}		96
C_2=aAeEiImM.bDfHjLnP.cCgGkKoO.dBhF1JpN	#33	$A_1\,A_2\,A_3\,A_4\,A_{18}$	B_3		96
C_3=aA.bD.cC.dB.eM.fP.gO.hN.iI.jL.kK.lJ.mE.nH.oG.pF	#34	$A_3\,A_4\,A_{19}$	B_{12}		96
C_4=aAi I.bDj L.cCkK.dBlJ.eMmE.fPnH.gOoG.hNpF	#35	$A_1\,A_2\,A_6$	B_1	C_1 pr	96
C_5=aAi I.bDj L.cCkK.dBlJ.eEmM.fHnP.gGoO.hFpN	#36	.	.	C_2 co	96
C_6=aA.bD.cC.dD.eG.fH.gE.hF.iJ.jI.kL.lK.mP.nO.oN.pM	#37	.	.	C_3 pr	96
C_7=aA.bI.cE.dM.eC.fK.gG.hO.iB.jJ.kF.lN.mD.nL.oH.pP	#38	.	.	C_4 co	96
C_8=aA.bD.cC.dB.eH.fG.gF.hE.iI.jL.kK.lJ.mP.nO.oN.pM	#39	$A_1\,A_2\,A_3\,A_4$	B_1	C_1 pr	96
C_9=aAi I.cCkK.bDjL.dBlJ.eHmPgFoN.fGnOhEpM	#40	.	.	C_5 co	96
C_{10}=aAbDcCdBeHfGgFhEiKjJkIlLmNnMoPpO	#41	.	.	C_3	96
C_{11}=aAcCbDdBeHgFfGhEiKkIjJlLmNoPnMpO	#42	.	B_3	C_3	96
C_{12}=aAcCbDdBeHgF.fGhEiIkKjLlJ.mPoN.nOpM	#43	.	B_4	C_4	160
C_{13}=aAi I.bDj L.cCkK.dBlJ.eHmP.fGnO.gFoN.hEpM	#44	.	$B_1\,B_3$	C_1	288
C_{14}=aAbBcCdDeP.fM.gN.hQ.iK.jL.kI.lJ.mF.nG.oH.pE	#45	.	.	C_3	288
	#46	.	.	C_7	288
	#47	.	$B_4\,B_3$	C_6	480
	#48	$A_9\,A_{10}\,A_{12}$	B_7	C_8 pr	96
	#49	.	.	C_9 co	96
	#50	.	.	C_{10} pr	96
	#51	.	.	C_{11} co	96
	#52	$A_9\,A_{10}\,A_{14}$	B_7	C_8 pr	96
	#53	.	.	C_{12} co	96
	#54	.	.	C_{13} co	96
	#55	.	.	C_{14}	96
	#56	$A_9\,A_4$	B_{11}	C_{14}	96

DEGREE 49

A_1 = ahrDMTb́bisENUć́cjtFHOd́dkuGIPé́eloAJQf́fmpBKRǵgnqCLSá́
A_2 = abcdefg.hijklmn.opqrstu.ABCDEFG.HIJKLMN.OPQRSTU.áb́ćd́éf́ǵ
A_3 = ahoAHOá́.bıpBIPb́.cjqCJQć́.dkrDKRd́.elsELSé́.fmtFMTf́.gnuGNUǵ
B_1 = bg.cf.de.há.iǵ.jf́.ké́.ld́.mć́.nb́.oO.pU.qT.rS.sR.tQ.uP.AH.BN.CM.DL.EK.FJ.GI
B_2 = hb́.ić́.jd́.ké́.lf́.mǵ.ná́.oQ.qS.sU.uP.pR.rT.tO.AK.DN.GJ.CM.FI.BL.EH
B_3 = hé́.if́.jǵ.ká́.lb́.mć́.nd́.sT.uO.pQ.rS.tU.oP.qR.AM.DI.GL.CH.FK.BN.EJ
B_4 = bć́.eB.cS.uL.gm.dN.fr.PD.iM.éT.gC.kq.oG.bn.OI.EK.ás.Rh.Hf́.jA.tQ
B_5 = bce.gfd.hoH.iqL.jsI.kuM.lpJ.mrN.ntK.Aá́O.Bć́S.Cé́P.Dǵ́T.Eb́Q.Fd́U.Gf́R
B_6 = bec.gdf.uLl.PDd́.ipM.Fé́T.gUC.Jkq.ćSB.Gb́O.mrN.Ino.Eá́R.Hjt.Khs.Af́Q
B_7 = hIt.iJu.jKo.kLp.lMq.mNr.nHs.ARb́.BSć́.CTd́.DUé́.EOf́.FPǵ.GQá́
B_8 = bec.dfg.hnj.ikl.ots.upq.AEG.CFD.IHK.JLM.OQR.PUT.b́f́á́.d́ǵé́
B_9 = bugP.iFǵ́J.ekdé́.LUDp.cMfC.qd́Tl.ć́GmI.EHKA.Sf́rj.b́hná́.BRNs.QotO
B_{10} = bigǵ.uJPF.eldD.kpé́U.cqfT.MlCd́.ć́EmK.GAlH.Sb́rn.f́ájh.BQNt.ROso
B_{11} = bć́gm.cSfr.dNeB.hTá́q.iHǵA.jCf́M.ksé́R.lnd́b́.oDOL.ptUQ.uIPG.EJKF
B_{12} = bEPAgKuH.iGJć́ǵIFm.cb́Cá́fnMh.qf́lSTjd́r.eQé́Odtko.LRpBDsUN
B_{13} = biEGPJAć́ǵǵKIuFHm.cqb́f́Clá́Sf́Tnjmd́hr.eLQRépOB dDtskUoN

U*			Order
#1	A_1	B_1	98
#2	"	B_5	147
#3	"	$B_1 B_5$	294
#4	$A_2 A_3$	B_1	98
#5	"	B_2	98
#6	"	B_5	147
#7	"	B_6	147
#8	"	B_7	147
#9	"	$B_1 B_2$	196
#10	"	B_9	196
#11	"	$B_5 B_1$	294
#12	"	$B_6 B_1$	294
#13	"	$B_5 B_2$	294
#14	"	$B_6 B_3$	294
#15	"	$B_7 B_1$	294
#16	"	$B_7 B_3$	294
#17	"	$B_8 B_3$	294
#18	"	$B_6 B_4$	294
#19	"	$B_9 B_2$	392
#20	"	$B_9 B_{10}$	392
#21	"	B_{12}	392
#22	"	$B_5 B_6$	441
#23	"	$B_5 B_9$	588
#24	"	$B_5 B_1 B_3$	588
#25	"	$B_6 B_1 B_3$	588
#26	"	$B_7 B_1 B_3$	588

U*			Order
#27	$A_2 A_3$	$B_6 B_4 B_1$	588
#28	"	$B_6 B_{11}$	588
#29	"	B_{13}	784
#30	"	$B_{12} B_2$	784
#31	"	$B_{12} B_{10}$	784
#32	"	$B_5 B_6 B_1$	882
#33	"	$B_5 B_6 B_3$	882
#34	"	$B_5 B_6 B_4$	882
#35	"	$B_5 B_{12}$	1176
#36	"	$B_5 B_9 B_2$	1176
#37	"	$B_5 B_9 B_{10}$	1176
#38	"	$B_6 B_1 B_3 B_4$	1176
#39	"	$B_9 B_{10} B_6$	1176
#40	"	$B_9 B_{10} B_7$	1176
#41	"	$B_{13} B_2$	1568
#42	"	$B_5 B_6 B_1 B_3$	1764
#43	"	$B_5 B_6 B_1 B_4$	1764
#44	"	$B_5 B_6 B_{11}$	1764
#45	"	$B_5 B_{13}$	2352
#46	"	$B_5 B_{12} B_2$	2352
#47	"	$B_5 B_{12} B_{10}$	2352
#48	"	$B_9 B_{10} B_6 B_{11}$	2352
#49	"	$B_5 B_6 B_1 B_3 B_4$	3528
#50	"	$B_5 B_9 B_{10} B_6$	3528
#51	"	$B_5 B_{13} B_2$	4704
#52	"	$B_5 B_9 B_{10} B_6 B_{11}$	7056

UNIVERSITY OF CHICAGO.

ARC-PRESERVING TRANSFORMATIONS.

By G. T. WHYBURN.

---

1. *Introduction.* If A is a compact continuum, a single valued transformation $T(A) = B$, where A and B are contained in a metric space, will be said to be *arc preserving* provided that the image of every simple arc in A is either a simple arc or a single point in B. Similarly, T will be said to be *irreducible* provided that no proper subcontinuum of A maps onto all of B under T.

Clearly any topological transformation is both arc preserving and irreducible. Obviously, also, if A contains no arc, then any transformation defined on A is arc preserving; and if B is itself a simple arc, then any continuous transformation sending A into B is arc preserving. The function $y = x^3 - x$ transforms the interval $(-2, 2)$ of the X-axis continuously into the interval $(-6, 6)$ of the Y-axis; and in this case the transformation is both arc preserving and irreducible; but is not a homeomorphism, since the roots of $x^3 - x - y$ are all real for $|y| \leq 2\sqrt{3}/9$ and are distinct for $|y| < 2\sqrt{3}/9$ whereas there is only one real root for $|y| > 2\sqrt{3}/9$.

Thus we see that even in case A and B are simple arcs (and therefore homeomorphic sets) and $T(A) = B$ is continuous, arc preserving and irreducible, T does not necessarily reduce to a homeomorphism. However, in this paper it will be shown, among other things, that if A is locally connected, if $T(A) = B$ is continuous and either A or B is cyclic (i. e., without cut points), then in order that T reduce to a homeomorphism it is necessary and sufficient that T be arc preserving and irreducible.

In this paper all transformations used will be supposed single valued and continuous.

It is readily seen that if A is a compact continuum and $T(A) = B$ is continuous then *there always exists a subcontinuum A_1 of A such that $T(A_1) = B$ and T is irreducible on A_1.* This follows at once from the Brouwer Reduction Theorem, since the property of being a subcontinuum of A mapping onto all of B under T clearly is inducible.

2. *Irreducible and arc-preserving transformations on locally connected continua.* Throughout this section we shall suppose A to be a compact locally connected continuum.

(2.1) *If A is a locally connected continuum and $T(A) = B$ is irreducible and arc preserving, then T is a homeomorphism on each true cyclic element of A on which T is not constant.*

Proof. Let E be such a true cyclic element of A. We have to prove that for each $b \in B$, $E \cdot T^{-1}(b)$ contains at most one point. Let us suppose this is not so. Then for some $b \in B$ there exists a non-degenerate arc $xpy \subset E$ such that $xpy \cdot T^{-1}(b) = x + y$. Now $T(xpy)$ is an arc ub and we may suppose p chosen so that $T(p) = u$. Let U_x and U_y be disjoint connected neighborhoods of x and y in A such that $T^{-1}(u) \cdot (U_x + U_y) = 0$ and $py \cdot U_x = xp \cdot U_y = 0$. Since there are non-cut points of A in both U_x and U_y, it follows from the irreducibility of T that there exist points $c \in U_x$ and $d \in U_y$ such that $T^{-1}[T(c)] \subset U_x$, $T^{-1}[T(d)] \subset U_y$. For if p is a non-cut point of A in U_x, say, there exists a neighborhood U_p of p such that $A - A \cdot U_x$ is contained in a single component N of $A - A \cdot U_p$; and since T is irreducible, there must exist a point c in U_x such that $N \cdot T^{-1}[T(c)] = 0$ and hence so that $T^{-1}[T(c)] \subset U_x$. Now let cq and dr be arcs in U_x and U_y respectively such that $cq \cdot xpy = q$, $dr \cdot xpy = r$. Let α denote the arc $cq + qpr + rd$.

We shall show that $T(\alpha)$ is not a simple arc. To do this it suffices to show that no one of the three points $T(c)$, $T(d)$, u separates the other two on $T(\alpha)$. Let β denote the subarc of the arc $T(qpr)$ from $T(q)$ to $T(r)$. Then since $T(q) \neq u \neq T(r)$, β does not contain u. Also $T(cq) \cdot u = T(dr) \cdot u = 0$. Thus $T(cq) + \beta + T(dr)$ is a connected subset of $T(\alpha)$ not containing u. But $T(cq + qp)$ is a connected subset of $T(\alpha)$ not containing $T(d)$ and $T(dr + rp)$ is a connected subset of $T(\alpha)$ not containing $T(c)$. Thus $T(\alpha)$ cannot be a simple arc, and our supposition that our theorem is not true leads to a contradiction.

(2.2) *Let $T(A) = B$ be irreducible and arc preserving, where A is locally connected. Then if B is cyclic, so also is A.*

Proof. Suppose, on the contrary, that A has a cut point p. Let us write $A = X + Y$ where X and Y are continua and $X \cdot Y = p$. Let $P = T^{-1}T(p)$ and $K = T^{-1}T(X)$.

(i) If $x \in Y \cdot (K - P)$, then for any arc px in A we have $px \subset K$.

For suppose not. Then for some $x' \in X$ we have $T(x') = T(x)$. Let $x'p$ be an arc in X. Then $T(x'p)$ is an arc in $T(X)$ and $T(px)$ is an arc in A not lying wholly in $T(X)$. Since each of the arcs $T(x'p)$ and $T(px)$ contains the two points $T(p)$ and $T(x)$, it follows that their sum contains a simple

closed curve. But $T(px) + T(x'p) = T(x'px)$ and this set must be a simple arc.

(ii) *If* $q \in (P - p)$, then for any arc pq in A we have either $pq \subset K$ or $pq \cdot K \subset P$.

For suppose pq contains a point x of $K - P$ and also a point y of $A - K$. Then by (i) we have $px \subset K$. Thus $y \subset xq$. Hence $T(px)$ is an arc in $T(X)$ and $T(xq)$ is an arc in B not lying wholly in $T(X)$. Since each of these arcs contains the two points $T(x)$ and $T(p)$ $[= T(q)]$, it follows that their sum contains a simple closed curve. This is impossible since $T(px) + T(xq) = T(pxq)$.

(iii) *If* R *is any component of* $A - K$, *then* $T[F(R)]$ *reduces to a single point.*[1]

For if not, there would exist an arc xy such that $T(x) \neq T(y)$, $x + y \subset K$ and $xy - (x + y) \subset R$. Now either $T(x)$ or $T(y)$ is $\neq T(p)$. Suppose $T(x) \neq T(p)$. Then by (i) any arc px in A is $\subset K$. Since then $px + xy$ is an arc pxy, it follows by (i) that $y \in P$. But since pxy contains $x \in (K - P)$ and also contains points in the open arc xy which are not in K, this contradicts (ii).

(iv) Now let R be a component of $A - K$ and let Q be the sum of all those components S of $A - K$ such that $T[F(S)] = T[F(R)] = a$. Then since $Q + T^{-1}(a)$ is closed, it follows that $T(Q) + a$ is closed. Since a is not a cut point of B it results that some point x of $T(Q)$ must be a limit point of $B - T(Q)$. Since Q is open in A, it follows from this that $T^{-1}(x)$ must intersect some component U of $A - K$ which does not belong to Q. Let $b = T[F(U)]$.

Now we have two essential cases to consider as follows: (I) $a \neq T(p) \neq b$, and (II) $a = T(p) \neq b$. In either case let $x_1 \in T^{-1}(x) \cdot Q$ and $x_2 \in T^{-1}(x) \cdot U$, let x_1a_1 and x_2b_1 be arcs such that $x_1a_1 - a_1 \subset Q$, $x_2b_1 - b_1 \subset U$, $a_1 \in T^{-1}(a)$, $b_1 \in T^{-1}(b)$.

In case I, we have a_1, $b_1 \in (K - P)$. Hence it follows by (i) that there exists an arc a_1b_1 in K. Then $T(a_1b_1)$ is an arc in $T(X)$ containing both a and b. But since $T(x_1) = T(x_2) = x$, it follows that $T(x_1a_1) + T(x_2b_1)$ contains an arc from a to b which clearly cannot lie in $T(X)$. Hence $T(a_1b_1) + T(x_1a_1) + T(x_2b_1) = T(x_1a_1 + a_1b_1 + b_1x_2) = T(x_1a_1b_1x_2)$ contains a simple closed curve, which is impossible.

In case II, consider an arc a_1p in A. If this arc contains b_1, then by (ii) we have $a_1p \subset K$ so that a_1p would contain a subarc a_1b_1 contained in K.

[1] $F(R)$ denotes the boundary of R relative to A, i.e., the set $\bar{R} - R$.

If a_1p does not contain b_1, then by (i) any arc pb_1 in A is $\subset K$; and $a_1p + pb_1$ contains an arc a_1b_1 some subarc a'_1b_1 of which is $\subset K$, where $a'_1 \in P = T^{-1}(a)$. Let us agree to set $a_1 = a'_1$ in case $a_1b_1 \subset K$. Then $T(a'_1b_1)$ is an arc in $T(X)$ containing both $T(a'_1) = a$ and $T(b_1) = b$. Since $T(x_1) = T(x_2) = x$, it follows that $T(x_1a'_1) + T(x_2b_1)$ contains an arc from $T(a'_1)$ to $T(b_1)$ which does not lie wholly in $T(X)$. Hence

$$T(x_1a'_1) + T(a'_1b_1) + T(b_1x_2) = T(x_1a'_1 + a'_1b_1 + b_1x_2) = T(x_1a'_1b_1x_2)$$

contains a simple closed curve, which is impossible.

(2.21) COROLLARY. *Let $T(A) = B$ be arc preserving, where A is locally connected and B is cyclic. If A has a cut point p, then for each component C of $A - p$ we have $T(C + p) = B$.*

This results at once from the proof of (2.2), because we can take X in that proof to be $C + p$; and it will be noted that the irreducibility of T was used only once, namely, in (iv) to insure the existence of at least one component R of $A - T^{-1}T(X)$.

(2.3) *If A is a locally connected continuum and $T(A) = B$ is irreducible and arc preserving and if for some $b \in B$, $x, y \in T^{-1}(b)$, $x \neq y$, then on any true cyclic element in the cyclic chain [2] $C(x, y)$, T must be constant.*

Proof. Let xpy be any arc in A from x to y. Now if $T(xpy) = b \in B$, our result follows from (2.1). Thus we suppose $T(xpy)$ contains more than one point. Also, without loss of generality, we can assume that $xpy \cdot T^{-1}(b) = x + y$. Hence $T(xpy)$ is an arc ub; and clearly we may suppose p chosen so that $T(p) = u$.

Now if we suppose (2.3) does not hold, there exists a true cyclic element E of A intersecting xpy in a non-degenerate arc cd where we have the order x, c, d, y and where T is not constant on E. Since by (2.1) T must be a homeomorphism on cd and since $T(xp) = T(py) = ub$, it follows that either $cd \subset xp$ or $cd \subset py$. Let us suppose the latter. Then we have the order x, p, c, d, y on xpy, where some points may coincide. Now there must exist a point w on $py - (p + y)$ which is a limit point of $E - cd$ and such that $u \neq T(w) \neq b$. For if not, then E would contain an arc $p'zy'$ such that $p'zy' \cdot xpy = p' + y'$ and $T(p') = u$, $T(y') = b$; but this is impossible,

[2] If M is a locally connected continuum, a subset of M is called an A-set provided it is closed and it contains every simple arc in M whose endpoints lie in it. If $x, y \in M$, the smallest A-set in M containing $x + y$ is called the *cyclic chain* $C(x, y)$. See my paper in *American Journal of Mathematics*, vol. 50 (1928), pp. 167-194; see also Kuratowski and Whyburn, *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331.

because then $T(xp' + p'zy') = T(p'y') + T(p'zy')$, which is a simple closed curve since $p'y' + p'zy'$ is a simple closed curve and T is a homeomorphism on every such curve by (2.1). (Note: $p'y'$ denotes the arc $p'y'$ of xy). Thus such a point w exists.

It follows readily [see the proof of (2.1)] that there exists an arc rv in E such that

$$rv \cdot xpy = v, \quad T(rv) \cdot (u + b) = 0, \quad \text{and} \quad T(r) \cdot ub = 0.$$

Now $xpv + rv$ is an arc in A; but $T(xpv) = T(xp) = ub$, whereas $T(rv)$ is an arc in B intersecting ub but containing neither u nor b and yet not being contained in ub. Thus $T(xpv + rv)$ cannot be a simple arc, contrary to hypothesis.

(2.4) *If A is locally connected and $T(A) = B$ is irreducible and arc preserving, then for each true cyclic element E_b in B there exists a unique true cyclic element E_a in A such that $T(E_a) = E_b$.*

Proof. Let $K = T^{-1}(E_b)$.

(i) For each component R of $A - K$, we have $T[F(R)] = p \in E_b$.

For suppose, on the contrary, that for some $x, y \in F(R)$ we have $T(x) \neq T(y)$. Then $T(R)$ is a connected subset of $B - E_b$ having at least the two limit points $T(x)$ and $T(y)$ in E_b. Clearly this is impossible since E_b is a cyclic element of B.

(ii) If we define the transformation $Z(A) = E_b$ as follows:

$$Z(x) = T(x) \text{ for } x \in K, \qquad Z(x) = T[F(R)] \text{ for } x \in R,$$

where R is a component of $A - K$, then Z is continuous and arc preserving.

The continuity is immediate. Let pq be any arc in A. Then $T(pq)$ is an arc ab in B. If $ab \cdot E_b = 0$, then $pq \subset$ some R; whence $Z(pq) = T[F(R)]$ = a single point. If $ab \cdot E_b \neq 0$, then $ab \cdot E_b$ is a subarc $a'b'$ of ab which may reduce to a single point. Let $x \in (pq - K \cdot pq)$. Then $Z(x) = T[F(R)]$, where R is the component of $A - K$ containing x, and clearly $Z(x) \in a'b'$. Thus $Z(pq) = a'b'$ and Z is arc preserving.

Let E_a be an A-set in A which intersects every one of the sets $T^{-1}(b)$, $b \in E_b$, and is irreducible with respect to the property of being an A-set intersecting every one of these sets. Then $Z(E_a) = E_b$ and Z is arc preserving on E_a since it is arc preserving on A. It now follows from (2.21) that E_a can have no cut point, and hence E_a is a true cyclic element of A. But T is a homeomorphism on each true cyclic element of A on which it is not constant, by (2.1), and hence is a homeomorphism on E_a; and since $T(E_a) \supset E_b$, this gives $T(E_a) = E_b$ as was to be proven.

Finally, there can exist no other true cyclic element $E_{a'}$ of A such that $T(E_{a'}) = E_b$. For if so, let x be a point of E_b such that $T^{-1}(x) \cdot [E_a + E_{a'}]$ contains no cut point of A. Then if $x_1 = E_a \cdot T^{-1}(x)$, $x_2 = E_{a'} \cdot T^{-1}(x)$, we have $C(x_1, x_2) \supset E_a + E_{a'}$ and this contradicts (2.3).

3. *Recapitulation. Conditions for homeomorphism.* The results just established in § 2 yield at once the following theorems.

(3.1) THEOREM. *If A is a compact locally connected continuum and $T(A) = B$ is irreducible and arc preserving, then T maps the true cyclic elements E_a of A on which T is not constant into the true cyclic elements E_b of B in a $(1-1)$ way and T is a homeomorphism on each E_a. (i. e., for each such E_a, $T(E_a)$ is an E_b and for each E_b there is a unique E_a such that $T(E_a) = E_b$).*

(3.2) THEOREM. *Let A be a compact locally connected continuum, let $T(A) = B$ be continuous and suppose that either A or B is cyclic. Then in order that T be a homeomorphism it is necessary and sufficient that it be irreducible and arc preserving.*

It is clear that for the validity of the former of these theorems the irreducibility of T is indispensable. However, if in the latter theorem we specify that B is cyclic, it seems possible that the condition of irreducibility on T could be dispensed with. This is indeed the case, as will now be shown, provided we assume further that A is hereditarily locally connected, i. e., that not only A but also every subcontinuum of A is locally connected.

(3.3) THEOREM. *Let A be hereditarily locally connected, let $T(A) = B$ be continuous and suppose that B is cyclic. Then in order that T be a homeomorphism it is necessary and sufficient that it be arc preserving.*

Proof. The necessity of the condition is obvious. We proceed to establish the sufficiency. In view of (3.2) it suffices to prove that T is irreducible. Let us suppose, on the contrary, that for some proper subcontinuum A' of A we have $T(A') = B$. We may suppose A' chosen so that T is irreducible on A'. Also, since T is arc preserving on A, clearly it is likewise arc preserving on A'. Furthermore, since A' is locally connected and B is cyclic, it follows by (3.2) that A' is cyclic and T is a homeomorphism on A'.

Now since $A' \neq A$, there exists an arc ab in A such that $ab \cdot A' = a + b$. Then $T(ab)$ is an arc and, since T is a homeomorphism on A', so also is $A' \cdot T^{-1}T(ab)$. Call this latter arc xy. Since A' is cyclic, there exists in A' an arc $\alpha\beta$ such that $xy \cdot \alpha\beta = \alpha + \beta$ and we have the order $x \leq \alpha < \beta \leq y$ on xy. Let $\alpha z \beta$ be the subarc of xy from α to β.

(i) If uv is any proper subarc of the arc $T(ab) = rs$, there exists a proper subarc pq of ab such that $T(pq) \supset uv$.

For let m be a point of $rs - uv$ such that $T(a) \neq m \neq T(b)$. Then either rm or ms contains uv, suppose the former. Then since $T^{-1}(m) \cdot (a + b) = 0$, ab contains a proper subarc pq where $p \in T^{-1}(m)$, $q \in T^{-1}(r)$; and since $T(pq)$ is a subarc of rs containing both r and m, we have $T(pq) \supset uv$. Similar reasoning applies in case $ms \supset uv$.

(ii) If $\alpha z\beta \neq xy$, then $T(\alpha z\beta)$ is a proper subarc uv of rs (since T is a homeomorphism on A'). Thus by (i) we have a proper subarc pq of ab such that $T(pq) \supset uv$. Not both a and b can belong to pq. Suppose b non-ϵ pq and suppose we have the order a, p, q, b on ab. Now $b \in xy$ and xy contains either an arc $b\beta$ not containing α or an arc $b\alpha$ not containing β. Call this arc t. Then in either case it is apparent that $\alpha\beta + t + pqb$ is a simple arc γ. But $T(\gamma) \supset T(\alpha\beta) + T(pq) = T(\alpha\beta) + uv = T(\alpha\beta) + T(\alpha z\beta)$; and this latter set is a simple closed curve, since $\alpha\beta + \alpha z\beta$ is a simple closed curve in A' and T is a homeomorphism on A'. Thus this case leads to a contradiction.

(iii) If $\alpha z\beta = xy$ we may suppose we have $\alpha = x$, $\beta = y$. Now both a and b lie on xzy and we may suppose we have the order $x \leq a < z < b \leq y$, on xzy. Since $T(a) \neq T(b)$, there exist arcs ac and db in ab and zb respectively such that $T(ac) \cdot T(db) = 0$. Since $T(ab) = T(xzy) = rs$, this gives $T(by) + T(xazd) \supset T(ac)$. Whence $T(by) + T(xazd) + T(cb) = rs$. Thus $T(cb + by + \beta\alpha + xazd) = T(cb) + T(xazd) + T(by) + T(\alpha\beta) = rs + T(\alpha\beta)$. But this is impossible, since $cb + by + \beta\alpha + xazd$ is a simple arc whereas $rs + T(\alpha\beta)$ is a simple close curve. This contradiction completes our proof.

4. *A-set reversing transformations.* If A is a compact locally connected continuum and $T(A) = B$ is continuous (as we shall suppose throughout this section), then T is said to be *A-set reversing* provided that for each $b \in B$, $T^{-1}(b)$ is either a single point or an A-set in A.

(4.1) THEOREM. *In order that T be A-set reversing it is necessary and sufficient that T be monotone*[3] *on every simple arc in A.*

To prove the condition necessary, suppose T *A-set* reversing, let pq be any simple arc in A and let $b \in T(pq)$. Then since $T^{-1}(b)$ is either a single point or an A-set, it follows that $pq \cdot T^{-1}(b)$ is[4] in either case a connected set. Consequently T is monotone on pq (indeed on any connected subset of A).

---

[3] A transformation $T(X) = Y$ is *monotone* provided that for each $y \in Y$, $T^{-1}(y)$ is connected. See C. B. Morrey, *American Journal of Mathematics*, vol. 57 (1935), pp. 17-50.

[4] For properties of A-sets, see Kuratowski and Whyburn, *loc. cit.*

The condition is also sufficient. For suppose it satisfied and let $b \in B$. Let $p, q \in T^{-1}(b)$ and let pq be any arc in A from p to q. Since by hypothesis T is monotone on pq and since $T(p) = T(q) = b$, we must have $T(pq) = b$ or $pq \subset T^{-1}(b)$. Thus $T^{-1}(b)$ is either a single point or an A-set and T is A-set reversing.

(4.11) COROLLARY. *Any A-set reversing transformation on A is arc preserving.*

For the property of being a simple arc is invariant under monotone transformations.

(4.12) COROLLARY. *If T is A-set reversing on A, then T is monotone on every connected subset of A.*

(4.13) *Any monotone transformation on a dendrite is arc preserving.*

(4.14) *Any monotone transformation on A is arc preserving on the set $K + H$ of all cut points and all end points of A.*

(4.2) *If A is a locally connected continuum and $T(A) = B$ is A-set reversing, then for every A-set A_b in B, $T^{-1}(A_b)$ is an A-set in A (or a single point).*

Proof. Let $p, q \in T^{-1}(A_b)$, $p \neq q$, and let pq be any arc in A from p to q. If $T(p) = T(q)$, then $T(pq) = T(p) \in A_b$ so that $pq \subset T^{-1}(A_b)$, since T is A-set reversing. If $T(p) \neq T(q)$, then by (4.11), $T(pq)$ is an arc from $T(p)$ to $T(q)$. Whence $T(pq) \subset A_b$ so that $pq \subset T^{-1}(A_b)$. Thus $T^{-1}(A_b)$ is an A-set.

(4.3) *If A is a locally connected continuum and $T(A) = B$ is A-set reversing, then for each true cyclic element E_b in B there exists a unique true cyclic element E_a in A such that $T(E_a) = E_b$ and T is a homeomorphism on E_a.*

Proof. Since T is monotone, there exists [5] a true cyclic element E_a of A such that $T(E_a) \supset E_b$. Now by (4.2), $T^{-1}(E_b)$ is an A-set in A. Since $T^{-1}(E_b)$ contains at least two points of E_a, this gives $E_a \subset T^{-1}(E_b)$. Whence $T(E_a) \subset E_b$. Accordingly $T(E_a) = E_b$. Finally, for no $b \in B$ can $T^{-1}(b)$ contain more than one point of E_a; for if it did, we would have $E_a \subset T^{-1}(b)$, which is impossible. Thus T is a homeomorphism on E_a. Obviously E_a is uniquely determined.

(4.31) COROLLARY. *If A and B are locally connected and cyclic and $T(A) = B$ is A-set reversing, T is a homeomorphism.*

THE UNIVERSITY OF VIRGINIA.

[5] See my paper in the *American Journal of Mathematics*, vol. 56 (1934), pp. 133-146.

AN ANALYTIC CHARACTERIZATION OF SURFACES OF FINITE LEBESGUE AREA. PART II.

By CHARLES B. MORREY, JR.[1]

In Part I of this paper,[2] the author gives an analytic characterization of surfaces of finite Lebesgue area in which only "non-degenerate" surfaces are admitted for discussion. The object of the present part is to generalize these results to arbitrary surfaces of finite area, the hypothesis of non-degeneracy being dropped.

Throughout this paper, we shall use the following vector notation: the letters x and X shall stand for the vector of the coördinates $(x^1, \cdots, x^N)$ and $(X^1, \cdots, X^N)$ of a point in the x-space in which a given surface will lie, the letters u and U for the coördinates of a point of the space containing the set on which the given surface will be parametrically represented, the sum and difference of pairs of these letters will denote the sum and difference vectors, x_α and $\partial x/\partial\alpha$ for the vector $(\partial x^1/\partial\alpha, \cdots, \partial x^N/\partial\alpha)$, α being a parameter, $x(u)$ and $X(U)$ will be vector functions, and if ϕ is a vector in any space, $|\phi|$ will denote its length. We may write $x(P)$ to mean $x(u)$ where u is the coördinate vector of P, u_P for the coördinate vector of P, etc. Given a point set E, $\bar{E}$ will denote its closure and E^* the set of its frontier points. All vector functions occurring in a transformation or the representation of a surface will be assumed continuous. As there are a great many topological terms used whose definitions are long, we shall merely give the reference to where they may be found in the author's paper, "The topology of path surfaces,"[3] which we shall hereafter refer to by the letter T. as frequent references to this paper will be made.

3. *The existence of a "generalized conformal" map of an arbitrary surface of finite Lebesgue area on a hemicactoid* (T. § 2, def. 12).

Definition 1. Let $\bar{H}$ be a hemicactoid and $\bar{h}$ a continuous curve (T. § 1, Definition 9) which is a subset of $\bar{H}$. Let $\bar{b}$ consist of all the points of $\bar{h}$ not separated from $\bar{B}$ in $\bar{H}$ (T. § 1, Definition 4) by any point of $\bar{h}$; the set $\bar{b}$ is a continuous curve which is either a subset of $\bar{B}$ or a subset of a single cyclic

[1] National Research Fellow (1931-33).

[2] *American Journal of Mathematics*, vol. 57 (1935), pp. 692-702.

[3] *American Journal of Mathematics*, vol. 57 (1935), pp. 17-50.

element (T. § 1, Definition 13) of $\bar{H}$ which does not contain all of that element unless that element is a point, in which case $\bar{b}$ is that point. If $\bar{b}$ is a subset of $\bar{B}$ which is a base set (not necessarily canonical) (T. § 2, Definitions 6 and 11), is a single point, or is a subset of a double cone (T. § 2, Definition 7) of $\bar{H}$ and does not separate that double cone being therefore homeomorphic with a base set, and if $\overline{(\bar{h} - \bar{b})}$ is a sum of canonical cactoids (T. § 2, Definitions 10 and 11), then we say that $\bar{h}$ is a *sub-hemicactoid of* $\bar{H}$ *and* $\bar{b}$ *is its base set.* $\bar{h}$ is not necessarily a hemicactoid in the ordinary sense as $\bar{b}$ may not be a canonical base set although it is always homeomorphic to one.

Definition 2. Let S be a surface represented by $x = X(U)$ on a hemicactoid $\bar{H}$ and let $\bar{h}$ be a sub-hemicactoid of $\bar{H}$ with base set $\bar{b}$. From T. § 4, Theorem 1 and the above remarks it follows that there exists a monotone transformation (T. § 2, Definition 5) $U = U(u)$, with corresponding collection of continua G (see T. § 2, Theorem 4), of a Jordan region $\bar{r}$ into $\bar{h}$ which carries the subcollection G_0 of T. § 2, Lemma 8, topologically (T. § 2, Definition 4 and Theorem 2) into $\bar{b}$. If we define $x(u) = X[U(u)]$, $u \in \bar{r}$, then $x = x(u)$ is a surface and all surfaces thus obtained are identical. We call this surface the *sub-surface of* S *corresponding to* $\bar{h}$.

THEOREM 1. *Let* S *be a surface represented on a hemicactoid* $\bar{H}$ *by* $x = X(U)$. *Let* $\bar{h}_1$ *and* $\bar{h}_2$ *be two sub-hemicactoids of* $\bar{H}$ *such that* $\bar{h}_1 + \bar{h}_2 = \bar{H}$ *and* $\bar{h}_1 \cdot \bar{h}_2$ *is a single cut point of* $\bar{H}$. *Then if* S_1 *and* S_2 *are the sub-surfaces of* S *corresponding, respectively, to* $\bar{h}_1$ *and* $\bar{h}_2$, *we have*

$$L(S) = L(S_1) + L(S_2).^4$$

Proof. Suppose $\bar{h}_1 \cdot \bar{h}_2$ is not the whole base set of either h_1 or $\bar{h}_2$. Then $\bar{b} \cdot \bar{b}_1$ and $\bar{b} \cdot \bar{b}_2$ each contain more than one point and $\bar{h}_1 \cdot \bar{h}_2 = \bar{b}_1 \cdot \bar{b}_2 \in \bar{b}$. Now, we may find a monotone transformation with corresponding collection G_1 of $\bar{r}_1$, $\bar{r}_1 : 0 \leqq u \leqq 3/8$, $0 \leqq v \leqq 1$, into $\bar{h}_1$ which carries the subcollection $G_{1,0}$ into $\bar{b}_1$, the segment, $u = 3/8$, $0 \leqq v \leqq 1$, being carried into the point $\bar{h}_1 \cdot \bar{h}_2$. Similarly we may carry $\bar{r}_2$, $\bar{r}_2 : 5/8 \leqq u \leqq 1$, $0 \leqq v \leqq 1$, on $\bar{h}_2$, the segment, $u = 5/8$, $0 \leqq v \leqq 1$, being also carried into $\bar{h}_1 \cdot \bar{h}_2$. In this case we unite these to give a map $U = U(u)$ of $\bar{H}$ on $\bar{Q}$ by first defining $U(u) = U_{\bar{h}_1 \cdot \bar{h}_2}$, $u \in \gamma$, $\gamma : 3/8 \leqq u \leqq 5/8$, $0 \leqq v \leqq 1$, and then defining $U(u)$ in $\bar{r}_1$ and $\bar{r}_2$ by the condition that $U = U(u)$ give the above determined mappings. Clearly $U = U(u)$ carries $G = G_1 + G_2$ topologically into $\bar{H}$ and its subcollection G_0 topologically into $\bar{B}$.

---

[4] $L(S)$ denotes the Lebesgue area (see introduction to part I) of S.

If $\bar{h}_1 \cdot \bar{h}_2$ is the whole base set of one of the $\bar{h}_i$, we shall call that one $\bar{h}_2$. We may map $\bar{h}_1$ on $\bar{Q}$ by $U = \bar{U}(u)$ as $\bar{h}_1$ was mapped on $\bar{r}_1$ above except that we require that a single point u_1 of $\bar{Q}$ be carried into $\bar{h}_1 \cdot \bar{h}_2$. We extend $\bar{U}(u)$ to the square $\bar{Q}_\lambda$, $\bar{Q}_\lambda : -\lambda \leqq u,\ v \leqq 1 + \lambda\ \ (\lambda > 0)$, by defining $\bar{U}(u) = \bar{U}(u_0)$ for $u \,\varepsilon\, \bar{Q}_\lambda - \bar{Q}$, u_0 being the point of intersection of the ray, $(1/2, 1/2)$, (u, v) with Q^*. Now u_1 is certainly interior to $\bar{Q}$. Let $u' = U'(u)$ be the radial contraction of $\bar{Q}_\lambda$ into $\bar{Q}$ with center at $(1/2, 1/2)$ and define $u'_1 = u'(u_1)$. Define the transformation (monotone but discontinuous at u'_1) $u'' = u''(u')$ of $\bar{Q}$ into $\bar{r}_1 + \gamma$, $\bar{r}_1$ being the doubly connected region consisting of all points $u'_1 + \lambda(\bar{u}'_0 - u'_1)$, $1/2 \leqq \lambda \leqq 1$, $\bar{u}'_0$ denoting any point of Q^*, by

$$u''(u') = u'_0 + (u' - u'_0)/2, \qquad u' \neq u'_1,$$

u'_0 being the last point of $\bar{Q}$ on the ray $u'_1 u'$; u'_1 is carried into all the points of the region γ, consisting of all points $u'_1 + \lambda(u'_0 - u'_1)$, $3/8 \leqq \lambda \leqq 1/2$, u'_0 ranging over all of Q^*. The transformation from u'' to u is continuous: $u = u(u'')$. Now let $r_2 = Q - (\bar{r}_1 + \gamma)$; we can map $\bar{h}_2$ on $\bar{r}_2$ by a monotone transformation $U = U(u'')$ which carries r^*_2 into $\bar{h}_1 \cdot \bar{h}_2$. If we define $U(u'') = \bar{U}[u(u'')]$ for $u'' \,\varepsilon\, \bar{r}_1 + \gamma$, then $U = U(u'')$, in this case also, gives a monotone transformation, with corresponding collection G, of $\bar{Q}$ into $\bar{H}$ which carries G_0 into $\bar{B}$.

Now in either case above S is represented on $\bar{Q}$ by $x = x(u) = X[U(u)]$. Let $x_0 = x(u_0)$, $u_0 \,\varepsilon\, \gamma$. Clearly, in either case above, the surfaces S_i are represented on $\bar{Q}$ by $x = x_i(u)$ $(i = 1, 2)$, where

$$x_1(u) = x(u), u \,\varepsilon\, \bar{r}_1;\ \ x_1(u) = x_0, u \,\varepsilon\, \gamma;\ \ x_1(u) = x_0, \quad u \,\varepsilon\, \bar{r}_2;$$
$$x_2(u) = x_0, \quad u \,\varepsilon\, \bar{r}_1;\ \ x_2(u) = x_0, u \,\varepsilon\, \gamma;\ \ x_2(u) = x(u), u \,\varepsilon\, \bar{r}_2.$$

Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n : x = \bar{x}_n(u)$, $u \,\varepsilon\, \bar{Q}$, be a sequence of polyhedra approaching S, whose areas approach $L(S)$, the $\bar{x}_n(u)$ being linear in patches in $\bar{Q}$ and approaching $x(u)$ uniformly. Let $\{r_m\}$ and $\{R_m\}$ be two sequences of numbers approaching zero so that r_m/R_m approaches zero. For each m, let n_m be $\geqq m$ and so large that $|\, \bar{x}_{n_m}(u) - x_0 \,| < r_m/2$, $u \,\varepsilon\, \gamma$. Now define: $\hat{x}_m(u) = \bar{x}_{n_m}(u)$ for those values of u in $\bar{Q}$ for which $|\, \bar{x}_{n_m}(u) - x_0 \,| \geqq R_m$, $x_m(u) = x_0$ for those values of u in Q for which $|\, \bar{x}_{n_m}(u) - x_0 \,| \leqq r_m$ (these constituting a finite number of regions bounded by a finite number of curves analytic in pieces, one of these regions containing γ), and

$$\hat{x}_m(u) = x_0 + [(|\, \bar{x}_{n_m}(u) - x_0 \,| - r_m)/(R_m - r_m)][\bar{x}_{n_m}(u) - x_0]$$

for the remaining values of u. This function $\hat{x}_m(u)$ is analytic in patches and equal to x_0 in the set in parenthesis above containing γ. Furthermore it is immediate that $L(\hat{S}_m) \leqq (R_m/(R_m - r_m)) \cdot L(\bar{\Pi}_{n_m})$. But now it is

well known that, in such a surface $\hat{S}_m$, we may inscribe a polyhedron $\bar{\Pi}$, $\bar{\Pi}: x = x(u)$ so that $\| S_m, \bar{\Pi} \|$ and $| L(S_m) - L(\bar{\Pi}) |$ are arbitrarily small and $\bar{x}(u) = x_0$, $u \in \gamma$. Hence it is clear that we may replace the sequence $\{\bar{\Pi}_n\}$ by a sequence $\{\Pi_m\}$, $\Pi_m: x = x_m(u)$, $u \in \bar{Q}$, where

$$\lim_{m\to\infty} \| \Pi_m, S \| = \lim_{m\to\infty} | L(\Pi_m) - L(S) | = 0,$$

the $x_m(u)$ being linear in patches, converging uniformly to $x(u)$, and with $x_m(u) = x_0$, $u \in \gamma$. Then if we define, for each m, $x_{1,m}(u)$ and $x_{2,m}(u)$ from $x_m(u)$ as $x_1(u)$ and $x_2(u)$ were formed from $x(u)$, it is clear that $L(\Pi_m) = L(\Pi_{1,m}) + L(\Pi_{2,m})$ and that $\{\Pi_{i,m}\}$ converges to S_i, $\Pi_{i,m}: x = x_{i,m}(u)$ $(i = 1, 2)$. Thus $L(S) \geqq L(S_1) + L(S_2)$.

Now, let $\{\bar{\Pi}_{i,n}\}$, $x = \bar{x}_{i,n}(u)$, $u \in \bar{Q}$, be sequences approaching S_i so that $L(\bar{\Pi}_{i,n})$ approaches $L(S_i)$, each $\bar{x}_{i,n}(u)$ being linear in patches over $\bar{Q}$ and the $\bar{x}_{i,n}(u)$ converging uniformly to $x_i(u)$ $(i = 1, 2)$. Repeat the above process on each sequence obtaining sequences $\{\Pi_{i,m}\}$, $\Pi_{i,m}: x = x_{i,m}(u)$, where $\Pi_{i,m}$ converges to S_i, $L(\Pi_{i,m})$ to $L(S_i)$, $x_{i,m}(u)$ uniformly to $x_i(u)$, each $x_{i,m}(u)$ being linear in patches on $\bar{Q}$ and equal to x_0 on $\bar{r}_{3-i} + \gamma$ $(i = 1, 2)$. Then define (for each m) Π_m by $\Pi_m: x = x_m(u)$ where

$$x_m(u) = x_{1,m}(u), u \in \bar{r}_1; \quad x_m(u) = x_0, u \in \gamma; \quad x_m(u) = x_{2,m}(u), u \in \bar{r}_2.$$

Then clearly $L(\Pi_m) = L(\Pi_{1,m}) + L(\Pi_{2,m})$ and $\lim_{m\to\infty} \| \Pi_m, S \| = 0$. Thus $L(S) \leqq L(S_1) + L(S_2)$ which completely demonstrates the theorem.

THEOREM 2. *Let S be a surface represented by $x = x(u)$ on a Jordan region $\bar{r}$, $x(u)$ being constant over r^*. Then there exists a sequence $\{\Pi_m\}$, $\Pi_m: x = x_m(u)$, of polyhedra (1) where $x_m(u)$ is non-degenerate except that it is constant over r^*, (2) $\lim_{m\to\infty} \| S, \Pi_m \| = 0$, and (3) $L(S) = \lim_{m\to\infty} L(\Pi_m)$.*

Proof. Take $\bar{r} = \bar{Q}$. Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n: x = \bar{x}_n(u)$, $u \in \bar{Q}$, be any sequence of polyhedra approaching S such that $L(\bar{\Pi}_n)$ approaches $L(S)$, the $\bar{x}_n(u)$ being linear in patches on $\bar{Q}$ and $\bar{x}_n(u)$ converging uniformly to $x(u)$. Let $x_0 = x(u_0)$, $u_0 \in Q^*$. By the process of Theorem 1, using this x_0, we may replace $\{\bar{\Pi}_n\}$ by a sequence $\{\bar{\Pi}_m\}$, $\bar{\Pi}_m: x = \bar{x}_m(u)$, satisfying (2) and (3), the $\bar{x}_m(u)$ converging uniformly to $x(u)$, being linear in patches on $\bar{Q}$ and equal to x_0 on Q^*. By moving the vertices of each $\bar{\Pi}_m$ slightly, we can obtain a sequence $\{\Pi_m\}$ of the desired type.

LEMMA 1. *If S is a surface represented by $x = X(U)$ on a hemicactoid $\bar{H}$ consisting of a segment, then $L(S) = 0$ whether $\bar{B}$ is the whole segment or just one end point.*

Proof. Case I: $\bar{H} = \bar{B} =$ a segment, $0 \leqq u \leqq 1$, $v = 0$. Then if we define $x^i(u, v) = X^i(u)$, $x^i = x^i(u, v)$, $0 \leqq v \leqq 1$ is a representation of S on Q. It is clear that we may approximate to S by polyhedra of the same form which obviously have zero area. Thus $L(S) = 0$ in this case.

Case II: $\bar{H}$ a vertical segment $\xi = \eta = 0$, $0 \leqq \zeta \leqq 1$, $(0, 0, 0)$ being $\bar{B}$. Then S can be represented on $\bar{Q}$ by $x = x(u) = X[U(u)]$, $U(u)$ being defined by $\xi = \eta = 0$,

$$\zeta = 1 - \sqrt{[(u-1/2)^2 + (v-1/2)^2]/[(U-1/2)^2 + (V-1/2)^2]},$$

where (U, V) is the last point of $\bar{Q}$ on the ray, $(1/2, 1/2)$, (u, v). It is clear in this case also that S can be approximated to by a sequence of polyhedra each of area zero.

LEMMA 2. *If S is a surface represented on a simple cyclic chain $\bar{C}$ (of type A or B* [see T. § 2, Definition 9]*), then $L(S)$ is the sum of the areas of the subsurfaces of S corresponding to the non-degenerate cyclic elements of the chain.*

Proof. Suppose S is represented on $\bar{C}$ by $x = X(U)$. For each n, we know that the number of non-degenerate cyclic elements of $\bar{C}$ which are of diameter $\geqq 1/n$ is finite. Hence, for each n, form $\bar{C}_n$ from $\bar{C}$ by replacing each non-degenerate cyclic element of $\bar{C}$ of length $< 1/n$ by its axis. If $\bar{C}$ is of type B, $\bar{C}_n \subset \bar{C}$ and we define $X_n(U) = X(U)$ on $\bar{C}_n$. If $\bar{C}$ is of type A, define $X_n(U) = X(U)$ on $\bar{C} \cdot \bar{C}_n$ and $X_n(U) = X(U')$ for U on the axis of one of the replaced double cones of $\bar{C}$, U' ranging linearly over the broken line consisting of the two equal sides of a generating triangle of that double cone as U ranges over the axis. Define S_n by $S_n : x = X_n(U)$, $U \in \bar{C}_n$. By the proof of T. § 4, Lemma 5, we see that $\lim_{n\to\infty} \| S, S_n \| = 0$. Let $P_0, \cdots, P_{k_n}$ be the end points arranged in order, of $\bar{C}$ and of the cyclic elements of $\bar{C}$ which are of diameter $\geqq 1/n$ and let $\bar{C}^{(1)}, \cdots, \bar{C}^{(k_n)}$ and $\bar{C}_n{}^{(1)}, \cdots, \bar{C}_n{}^{(k_n)}$ be the parts into which $\bar{C}$ and $\bar{C}_n$, respectively, are divided by these points, the $S^{(i)}$ and $S_n{}^{(i)}$ being the corresponding subsurfaces of S and S_n. Each $\bar{C}_n{}^{(i)}$ is either a segment or a cyclic element of $\bar{C}$ of diameter $\geqq 1/n$. By repeated application of Theorem 1,

$$L(S) = \sum_{i=1}^{k_n} L(S^{(i)}), \qquad L(S_n) = \sum_{i=1}^{k_n} L(S_n{}^{(i)}).$$

Since $S_n{}^{(i)}$ is either identical with $S^{(i)}$ or corresponds to a segment of $\bar{C}_n$, in

which case $L(S_n^{(i)}) = 0$, it follows immediately that $L(S_n) \leqq L(S)$ and hence that $L(S)$ is equal to $\sum_{i=1}^{\infty} L(S'^{(i)})$ where the $S'^{(i)}$ are the subsurfaces of S corresponding to non-degenerate cyclic elements of $\bar{C}$.

LEMMA 3. *Let S be a surface represented by $x = x(u)$ on a Jordan region $\bar{r}$. Let γ be a maximal continuum over which $x(u)$ is constant which separates $\bar{r}$, and let d be a component of $\bar{r} - \gamma$ which may be a bounded complementary domain* (T. footnote p. 31) *of γ or a set consisting of a region, bounded by a portion of γ and a portion (possibly all) of r^*, plus that portion of r^* not in γ. Define $x_1(u) = x(u)$ for $u \in \bar{r} - d$, and $x_1(u) = x(u_0)$ for $u \in d$, u_0 being in γ. Then, if we define S_1 by $S_1 : x = x_1(u)$, $u \in \bar{r}$,*

$$L(S_1) \leqq L(S).$$

Proof. Take $\bar{r} = \bar{Q}$. Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n : x = \bar{x}_n(u)$, $u \in \bar{Q}$, be a sequence of polyhedra approaching S such that $\lim_{n\to\infty} L(\bar{\Pi}_n) = L(S)$, the $\bar{x}_n(u)$ being linear in patches and converging uniformly to $x(u)$ on $\bar{Q}$. By the method of Theorem 1, we may replace $\{\bar{\Pi}_n\}$ by a sequence $\{\Pi_n\}$ of polyhedra, $\Pi_n : x = x_n(u)$, where $\lim_{n\to\infty} \| S, \Pi_n \| = \lim_{n\to\infty} | L(S) - L(\Pi_n) | = 0$, and where the $x_n(u)$ are linear in patches, approach $x(u)$ uniformly, and $x_n(u) = x_0 = x(u_0)$ in some closed connected region γ_n of finite connectivity, bounded by polygons, and including γ in its interior, the u_0 above being in γ.

Now, for each n, let d_n consist of all of the (finite number) of components of $\bar{Q} - \gamma_n$ which lie in d. Now, clearly, since γ is maximal, $\gamma = \prod_{n=1}^{\infty} \gamma_n$. Hence if, for each n, we define $x_{1,n}(u) = x_n(u)$, $u \in \bar{Q} - d_n$ and $x_{1,n}(u) = x_0$, $u \in d_n$, then the surfaces $\Pi_{1,n}$, $\Pi_{1,n} : x = x_{1,n}(u)$, are polyhedra approaching S_1 where also $L(\Pi_{1,n}) \leqq L(\Pi_n)$. Hence

$$L(S_1) \leqq \underline{\lim_{n\to\infty}} L(\Pi_{1,n}) \leqq \lim_{n\to\infty} L(\Pi_n) = L(S)$$

which proves the lemma.

THEOREM 3. *If S is a surface represented on a hemicactoid $\bar{H}$, then $L(S)$ is the sum of the Lebesgue areas of the subsurfaces of S corresponding to the non-degenerate cyclic elements of $\bar{H}$.*

Proof. Let $\bar{H} = \sum_{m=1}^{\infty} \bar{B}_m + \sum_{n=1}^{\infty} \bar{C}_n + H^*$, where H^* is a completely disconnected set of end points of $\bar{H}$ which are limit points of $\Sigma \bar{B}_m + \Sigma \bar{C}_n$; the

$\bar{B}_m$ are the simple chains of type B of the base set and $\bar{C}_n$ are those of type A. We may arrange all of these into a single sequence $\{\bar{D}_n\}$ such that $\bar{D}_n \cdot (\bar{D}_1 + \cdots + \bar{D}_{n-1})$ is a single point. Let S be represented on $\bar{H}$ by $x = X(U)$ and let $U = U(u)$ be a monotone transformation of a Jordan region $\bar{r}$ onto $\bar{H}$. For each n let $\bar{H}_n = \bar{D}_1 + \cdots + \bar{D}_n$, let G_n be the continua of $\bar{r}$ carried into $\bar{H}_n$ by $U = U(u)$, and let $\hat{G}_n$ be the set of points of $\bar{r}$ covered by G_n. Now it is easy to see (using the method of T. § 2, Lemma 8) that $\hat{G}_n$ is a continuum and that each point of $\bar{r}$ either belongs to $\hat{G}_n$, is in a bounded complementary domain of a continuum of G_n, or is in a "region" bounded by a portion of a continuum g_n of G_n and a portion of r^* where we include this portion of r^* not in g_n in the "region." Now let G'_n be the collection of continua g'_n, each g'_n being obtained from the corresponding g_n by adding to it all of its complementary domains and all of the "regions" bounded by a portion of g_n and a portion of r^*. Define $U_n(u) = U(u)$ for $u \varepsilon \hat{G}_n$ and $U_n(u) = U(u_0)$ for u in any of the components of $\bar{r} - \hat{G}_n$, where u_0 belongs to the g_n corresponding to the g'_n containing u. Clearly $U = U_n(u)$ is a monotone transformation of $\bar{r}$ into $\bar{H}_n$ carrying G'_n into $\bar{H}_n$ and $G'_{n,0}$ into $\bar{B}_n$, the base set of $\bar{H}_n$. Let S_n be defined by $x = x_n(u) = X[U_n(u)]$; clearly $\{U_n(u)\}$ converges uniformly to $U(u)$ and S_n converges to S.

Now, by repeated use of Lemma 3 and a simple limit process which merely makes use of the lower semicontinuity of $L(S)$, we see that $L(S_n) \leqq L(S)$ and hence $L(S) = \lim_{n\to\infty} L(S_n)$. By Lemma 2 and Theorem 1, $L(S_n)$ is the sum of the areas of the subsurfaces of S_n (also subsurfaces of S) which correspond to the non-degenerate cyclic elements of $\bar{H}_n$. Hence $L(S)$ is the sum of the areas of the subsurfaces of S corresponding to non-degenerate cyclic elements of $\bar{H}$.

Definition 3. A surface S is said to be *represented generalized conformally* on a double cone $\bar{C}$ by $x = X(U)$ if, when any Jordan region of $\bar{C}$ is mapped conformally (with the obvious conventions at the vertices or equatorial edge) on a plane Jordan region $\bar{r}$ by $U = U(u)$ (this being easily seen to be possible), the function $x(u) = X[U(u)]$ is a generalized conformal vector function, i. e., the components $x^i(u, v)$ are all A. C. T. (Part I, § 2, Definition 1) and the formal expressions E, F, and G satisfy $E = G$, $F = 0$ almost everywhere. It is clear that in such a representation on $\bar{C}$, there is a function M, summable over $\bar{C}$, which is the transform of $E = G$ by the transformation $U = U(u)$, $L(S)$ being given by

$$L(S) = \int\int_{\bar{C}} M \, d\sigma,$$

$d\sigma$ being the element of area on $\bar{C}$. This definition is seen to be consistent for it is easily seen that if $u = u(u')$ is a conformal map of $\bar{r}$ on itself and $x'(u') = x[u(u')]$, then $x'(u')$ is also a generalized conformal vector function on $\bar{r}$.

LEMMA 4. *If $x = X(U)$, $U \in \bar{C}$, is a generalized conformal representation of S on a double cone $\bar{C}$ and if $U = U(U')$ is a conformal map of $\bar{C}'$ on $\bar{C}$, then $x = X'(U') = X[U(U')]$ is a generalized conformal representation of S on $\bar{C}'$.*

Proof. Let $\bar{R}$ and $\bar{R}'$ be corresponding Jordan regions on $\bar{C}$ and $\bar{C}'$ respectively (the conformal transformation of $\bar{C}'$ into $\bar{C}$ being $1-1$ and continuous as is easily seen). It is clear that if we map $\bar{R}$ on a plane Jordan region $\bar{r}$ conformally that the product of this transformation and $U = U(U')$ yields a conformal map of $\bar{R}'$ on $\bar{r}$. Then the lemma follows from Definition 3.

LEMMA 5.[5] *Any polyhedron Π which may be mapped non-degenerately on the surface Σ of a sphere may be mapped generalized conformally on Σ, any three logically distinct points* (Part I, § 2, Definition 4) *of Π being made to correspond with three given distinct points of Σ. It is easy to see that we may map on a double cone also, for simple formulas give a $1-1$ conformal (with the obvious conventions) representation of a double cone on the sphere.*

LEMMA 6. *Let S be a surface of finite Lebesgue area represented non-degenerately on a double cone $\bar{C}$, one end of which is the base set. Then S can be represented generalized conformally on a double cone $\bar{C}_1$ ($= \bar{C}$ perhaps), a given end point being its base set, this point and two others being made to correspond to the first end point and two other points of $\bar{C}$.*

Proof. By Theorem 2, we may choose a sequence $\{\Pi_n\}$ of polyhedra representable non-degenerately on $\bar{C}$ which approach S and are such that $\lim_{n\to\infty} L(\Pi_n) = L(S)$. It is immediate that we may choose a sequence of non-degenerate representations $x = X_n(U)$, $U \in \bar{C}$, of Π_n on $\bar{C}$ which approach the non-degenerate representation $x = X(U)$ of S on $\bar{C}$. Now let A, B, and C be three distinct points on $\bar{C}$, A say being the base set of $\bar{C}$, and let A', B', and $\bar{C}'$ be three distinct points on $\bar{C}'$, A' being its base set. Now, by Lemma 5 each Π_n may be mapped generalized conformally on $\bar{C}'$, by $x = X'_n(U')$, $U' \in \bar{C}'$, the points of Π_n corresponding by $x = X_n(U)$ to A, B, and C being mapped

[5] See, for instance, C. Caratheodory, "Conformal representation," *Cambridge Tracts in Mathematics and Mathematical Physics*, No. 28, § 161 and §§ 125-128, particularly.

on A', B', and C' respectively. An argument similar to that of Theorem 1, § 2 (Part I) of this paper shows that we may extract a uniformly convergent subsequence of the $X'_n(U')$ with limit function $X'(U')$, and it is easily seen that $x = X'(U')$ gives a generalized conformal map of S on $\bar{C}'$.

THEOREM 4. *A necessary and sufficient condition that a surface S be of finite area is that there exist a hemicactoid $\bar{H}$ on which S may be represented continuously the representation being generalized conformal on each non-degenerate cyclic element of $\bar{H}$, with M summable over $\bar{H}$. For such representations, $L(S)$ is given by*

$$L(S) = \int\int_{\bar{H}'} M \, d\sigma,$$

$\bar{H}'$ being the sum of the non-degenerate cyclic elements of $\bar{H}$, $d\sigma$ being the element of area on $\bar{H}$.

Proof. The second statement of the theorem and hence the sufficiency of the first statement follows from Lemma 5, § 2 (Part I of this paper), Theorem 2, and Theorem 3.

To show the existence of such a representation of a surface S of finite area, first let S be represented non-degenerately by $x = X'(U')$ on a hemicactoid $\bar{H}' = \bar{B}' + \sum_{m=1} \bar{C}'_m + H'^*$ (T. § 4, Theorem 2), $\bar{B}'$ being its base set and $\bar{C}'_m$ its canonical cactoids. We have $\bar{B}' = \sum_{n=1}^{\infty} \bar{B}'_n + B'$ and $\bar{C}'_m = \sum_{n=1}^{\infty} \bar{C}'_{m,n} + \bar{C}'_m$ where the B' and C'_m are mutually exclusive completely disconnected sets and $\bar{B}'_n$ and $\bar{C}'_{m,n}$ are simple cyclic chains of types B and A respectively joined together as in T. Definition 12, § 2, and T. Theorem 1, § 1.

Let the subsurface corresponding to a non-degenerate cyclic element of one of these chains be represented generalized conformally (since it is of finite area) on that element by $x = \bar{X}'(\bar{U}')$ the points corresponding under $x = X'(U')$ to the end points of the element being mapped on those same end points (Lemma 6). Now on each $\bar{B}'_n$ and $\bar{C}'_{m,n}$ let continuous monotone transformations $T_n^{(2)}$ and $T_{m,n}^{(1)}$ (for $\bar{B}'_n$ and $\bar{C}'_{m,n}$ respectively) be determined as follows: they are to be the identity at all points of their respective chains not in a non-degenerate cyclic element; on a non-degenerate cyclic element let these transformations be one of the monotone transformations $U' = \bar{U}'(U')$, induced by the relation $X'[U'(\bar{U}')] = \bar{X}'(\bar{U}')$, of that element into itself which carries each end point into itself, such a transformation existing by T. § 4, Theorem 3; these $T^{i}_{k,l}$ are easily seen to be continuous over the whole chain since in each chain, there are only a finite number of non-degenerate cyclic

elements of diameter $\geqq \varepsilon$ for each $\varepsilon > 0$. Now, these transformations satisfy the hypotheses of T. Theorem 5, § 2, so that we can find a hemicactoid $\bar{H}$ with base set $\bar{B}$, etc., such that (1) there are $1-1$ continuous transformations $S^{(1)}_{m,n}$ of $\bar{C}'_{m,n}$ into $\bar{C}_{m,n}$ and $S_n{}^{(2)}$ of $\bar{B}'_n$ into $\bar{B}_n$ which are conformal (with the obvious conventions at the vertices and edges) between corresponding non-degenerate cyclic elements and (2) the transformations $[S^{(1)}_{m,n}]^{-1}T^{(1)}_{m,n}$ and $[S_n{}^{(2)}]^{-1}T_n{}^{(2)}$, defined respectively on $\bar{C}_{m,n}$ and $\bar{B}_n$ unite to form a continuous monotone transformation $U'(U)$ of $\bar{H}$ into $\bar{H}'$ which carries $\bar{B}$ into $\bar{B}'$. If we define $X(U) = X'[U'(U)]$, $U \in \bar{H}$, $x = X(U)$ is clearly a representation of S on $\bar{H}$, and this representation is generalized conformal on each non-degenerate cyclic element of $\bar{H}$, since for each such element, it is the product of $[S_n{}^{(2)}]^{-1} \cdot R'$ or $[S^{(1)}_{m,n}]^{-1} \cdot R'$, where R' is the generalized conformal representation $x = \bar{X}'(\bar{U}')$, such a product being generalized conformal by Lemma 4.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

SOME SEPARATION PROPERTIES OF THE PLANE.[1]

By R. E. Basye.

In a previous paper [2] the writer defined the notion of a simply connected set and showed how it was useful in proving a number of separation theorems. Our purpose here is to extend this notion to multiple connectivity and derive some further results for the plane. The abbreviation S. C. will hereafter be used in referring to the paper mentioned above.

Let n be a positive integer. A connected set M is said to be *multiply connected of order n*, or *n-ply connected*, if (1) for each pair of points A and B of M and any relatively closed subset L of M that separates A from B in M there exists a subset of L, consisting of n or less components, which separates A from B in M, and (2) there exists at least one pair of points A and B of M and a relatively closed subset L of M that separates A from B in M such that every subset of L which separates A from B in M has at least n components. This definition becomes the criterion for a connected set to be *n-ply connected in the weak sense* if "separates" is replaced throughout by "weakly disconnects." [3]

Theorem 1. *If, in a metric space, α is a monotonic descending sequence of compact continua which are multiply connected in the weak sense of orders not greater than n, then the product of the sets of α is also multiply connected in the weak sense of order not greater than n.*

The proof is similar to that of Theorem 12 of S. C.

Theorem 2. *If D is a bounded and connected subdomain of the plane whose boundary consists of n mutually exclusive simple closed curves, then $\bar{D}$ is n-ply connected.*

Let F be a closed subset of $\bar{D}$ which separates a point P_1 from a point P_2 in $\bar{D}$. Let F' be a closed subset of F which is the common boundary of two

[1] Presented to the American Mathematical Society, September 7, 1934, under the title, "Multiply connected sets."

[2] "Simply connected sets," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 341-356.

[3] A subset L of a connected set M is said to weakly disconnect a point A from a point B in M if L intersects every connected and relatively closed subset of M which contains $A + B$.

connected domains D_1 and D_2, relative to $\bar{D}$, containing P_1 and P_2 respectively. Then $\bar{D}$ is the sum of two continua, $\bar{D}_1$ and $\bar{D} - D_1$, whose common part is F'. If F' has $n+1$ or more components, it follows from a theorem of S. Straszewicz [4] that $\bar{D}$ has at least $n+1$ complementary domains. But this is not the case. Hence F' has not more than n components.

On the other hand let $J_1, \cdots, J_n$ denote the simple closed curves which constitute the boundary of D. It is known that if Δ is a bounded and connected subdomain of the plane, and b is an arc which lies in Δ except for its endpoints, which lie in the boundary β of Δ, then $\Delta - \Delta \cdot b$ is connected or the sum of two mutually exclusive connected domains Δ_1, Δ_2 according as the endpoints of b lie in different or in the same component of β; in the latter case b lies on the boundary of each of the domains Δ_1, Δ_2. With the aid of this fact there can be constructed n mutually exclusive arcs $b_{12}, b_{23}, \cdots, b_{n-1n}, b_{n1}$ such that b_{kl} lies in D except for its endpoints, which lie in J_k and J_l respectively; and $D - D \cdot \Sigma b_{kl}$ is the sum of two mutually exclusive connected domains d_1, d_2 such that Σb_{kl} lies on the boundary of each. Furthermore it can be seen that $\bar{d}_1 - \Sigma b_{kl}$ and $\bar{d}_2 - \Sigma b_{kl}$ are mutually separated sets. Hence Σb_{kl}, which has n components, irreducibly separates d_1 from d_2 in $\bar{D}$.

It follows that $\bar{D}$ is *n-ply* connected.

In passing we remark that, with the aid of Theorem 2 and a generalization of Theorem 7 of S. C., it can be shown that a connected subdomain of the plane is n-ply connected if and only if its complement has exactly $n-1$ bounded components.

THEOREM 3. *Every compact plane continuum which has exactly n complementary domains is multiply connected in the weak sense of order not greater than n.*

Such a continuum can be expressed as the product of a monotonic descending sequence of compact continua each of which is the closure of a connected domain whose boundary consists of n mutually exclusive simple closed curves. Hence Theorem 3 follows from Theorems 1 and 2.

THEOREM 4. *If A and B are two points of a compact plane continuum M having exactly n complementary domains, and G is a countable collection of mutually exclusive closed subsets of M such that G* is closed and weakly disconnects A from B in M, then there exist n elements or less of G whose sum weakly disconnects A from B in M.*

[4] "Über die Zerschneidung der Ebene durch abgeschlossene Mengen," *Fundamenta Mathematicae*, vol. 7 (1925), p. 173, Theorem II.

By Theorem 3 there exists a closed subset F of G^* which weakly disconnects A from B in M and has not more than n components. No component of F can intersect more than one element of G since, if it did, it would be the sum of countably many (more than one) mutually exclusive closed sets. Hence F is a subset of the sum of n or less elements of G.

THEOREM 5. *Let A and B be two points of a connected and locally arcwise connected metric space S, and let G be a countable collection of closed sets such that* (1) *the common part of every pair of elements of G is the closed set H (which may be vacuous),* (2) *if $b_1, \cdots, b_{n+1}$ are $n+1$ arcs from A to B that lie in $S - H$, then $b_1 + \cdots + b_{n+1}$ lies in a compact set which is multiply connected in the weak sense of order n or less and whose closure contains no point of H, and* (3) *G^* is locally compact. If G^* separates A from B in S then G contains n or less elements whose sum separates A from B in S.*

The proof is similar to that of Theorem 3 of S. C.

LEMMA Q. *In a sphere S let G be a countable collection of closed sets such that the common part of every pair of elements of G is N, a closed set having exactly n components, where n is a positive integer. If G^* separates a point A from a point B in S, then G contains n or less elements whose sum separates A from B in S.*

Consider any $n+1$ arcs from A to B which have no point in common with N, and let β denote their sum. If $N_1, \cdots, N_n$ denote the components of N, let $J_1, \cdots, J_n$ denote n mutually exclusive simple closed curves such that $J_i (i = 1, \cdots, n)$ separates β from N_i in S. The boundary of the complementary domain Δ of $J_1 + \cdots + J_n$ which contains β consists of one or more of the curves J_i; hence $\bar{\Delta}$ is homeomorphic to the closure of a bounded and connected plane domain whose boundary consists of n or less mutually exclusive simple closed curves. Hence, by Theorem 2, $\bar{\Delta}$ is multiply connected of order not greater than n. Therefore, by Theorem 5, G contains n or less elements whose sum separates A from B in S.

With the aid of the preceding lemma we can generalize, as follows, a theorem of R. L. Moore.[5]

THEOREM 6. *In a plane S let G be a countable collection of closed sets such that* (1) *the common part of every pair of elements of G is N, a closed*

[5] "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, vol. 13, p. 298, Theorem 113.

set having exactly n components, where n is a positive integer, and (2) *there exists not more than one element g of G such that* $g - N$ *is not compact. If* G^* *separates a point A from a point B in S, then G contains n or less elements whose sum separates A from B in S.*

Let Σ be a sphere, P a point of Σ, and H a homeomorphism of S into $\Sigma - P$. Let G', N', A', B' be the images of G, N, A, B, respectively, under H. Let G'' be the collection whose elements are defined as follows. Let g' denote an element of G'. If P is not a limit point of g' then g' is an element of G'', and if P is a limit point of g' then $g' + P$ is an element of G''. If P is not a limit point of any element of G' then $N' + P$ is an element of G''. The elements of G'' are closed in Σ and the common part of each pair is the same set, which we call N''. Now N'' does not have more than n components. Hence, since G''^* separates A' from B' in Σ, there exists, by Lemma Q, n or less elements of G'' whose sum ν separates A' from B' in Σ. Thus $\nu - \nu \cdot P$ separates A' from B' in $\Sigma - P$ and moreover is the sum of not more than n elements of G'. Therefore the image of $\nu - \nu \cdot P$ under H separates A from B in S and is the sum of not more than n elements of G.

In Theorem 21 of S. C. the writer has given a certain extension of a result [6] of Rutt and Roberts. With the aid of Theorem 6 we can extend this result in another direction, as follows.

THEOREM 7. *In a plane S let N be a closed set having exactly n components, where n is a positive integer. Let G be any collection of closed sets and H a countable subcollection* (*possibly vacuous*) *of G such that* (1) *the common part of each pair of elements of G is N,* (2) *every component of every element of* $G - H$ *intersects N,* (3) *there exists not more than one element q of G such that* $g - N$ *is not compact, and* (4) G^* *is closed. If* G^* *separates a point A from a point B in S then G contains n or less elements whose sum separates A from B in S.*

Suppose the contrary. Let G' denote the collection of all elements each of which is either an element of H or a component of the sum of N and a component of a set obtained by subtracting N from an element of $G - H$. Let H' denote the collection whose elements are those of H and that element g of G, if there exists one, such that $g - g \cdot N$ is not compact. There exists a subset F of G'^* which is closed, separates A from B in S, contains every

[6] N. E. Rutt, "On certain types of plane continua," *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 815, Theorem IV and Corollary IV; and J. H. Roberts, "Concerning collections of continua not all bounded," *American Journal of Mathematics*, vol. 52 (1930), p. 553, Theorem I.

element of G' with which it has a point of $G'^* - (H^* + N)$ in common, and is irreducible with respect to these three properties. The set F is a continuum.

With the aid of Theorem 6 there can be constructed $4n$ arcs $b_1, \cdots, b_{4n}$ from A to B which have no points in common with $H'^* + N$ and such that, if P_i $(i = 1, \cdots, 4n)$ denotes the first point of b_i in the order from A to B which lies in F, the points P_i all lie in different elements of G'. The element g'_i of G' which contains P_i contains a component of N. Hence there exists a component of N which is common to four of the elements g'_i, say g'_1, g'_2, g'_3, g'_4. By an argument similar to one used in Theorem 21 of S. C. it can be shown that two of these four continua, say g'_1 and g'_3, have the property that $g'_1 + g'_3$ separates $g'_2 - g'_2 \cdot N$ from $g'_4 - g'_4 \cdot N$ in F. Thus $F - (g'_1 + g'_3) = F_2 + F_4$, where F_2 and F_4 are mutually separated sets containing P_2 and P_4 respectively. Consider the two sets $R_2 = F_2 + (g'_1 + g'_3)$ and $R_4 = F_4 + (g'_1 + g'_3)$. The common part of R_2 and R_4 is a continuum and their sum is F. Hence either R_2 or R_4, say the former, separates A from B in S. But R_2 is a proper closed subset of F which contains every element of G' with which it has a point of $G'^* - (H^* + N)$ in common. This is contrary to the construction of F.

If the set N is not restricted as to the number of its components, the following weaker conclusion will hold true.

THEOREM 8. *In a plane S let N be a non-vacuous closed set, G any collection of closed sets, and H a countable subcollection (possibly vacuous) of G such that* (1) *the common part of each pair of elements of G is N,* (2) *every component of every element of $G - H$* intersects *N,* and (3) *G^* is closed. If G^* separates a point A from a point B in S, then G contains a countable subcollection K such that K^* intersects every compact continuum which contains $A + B$.*

Suppose that G^* is compact. Let $\Delta_1, \Delta_2, \cdots$ be a sequence of domains closing down on N such that the closure of each domain has not more than a finite number of components and contains neither A nor B. Consider the collection G_i $(i = 1, 2, \cdots)$ of all elements each of which is the sum of $\bar{\Delta}_i$ and an element of G. This collection satisfies the conditions of Theorem 7 and hence contains a finite number of elements whose sum L_i separates A from B in S. Denote by M_i the finite collection of those elements of G which intersect $L_i - \bar{\Delta}_i$. We take for K the collection of all elements of G which are elements of some M_i. Let C be a continuum which contains $A + B$. If C intersects N it intersects every element of K. If C does not intersect N there exists an integer r such that $\bar{\Delta}_r$ contains no point of C; therefore C contains a point of $L_r - \bar{\Delta}_r$ and hence a point of an element of K.

The case where G^* is not compact can be reduced by an inversion of the plane to the one considered.

It has been proved by Kuratowski [7] that if three compact plane continua have a point in common, and their sum separates a point A from a point B in the plane, then there exists a pair of these continua whose sum separates A from B in the plane. The writer has elsewhere [8] obtained a generalization of this result for which the containing space is any subcontinuum of the plane. An extension in another direction, in which the number of continua is not restricted, will here be given.

THEOREM 9. *In a plane S let N be a non-vacuous closed set and let G be any collection of continua such that* (1) *no two elements of G have a point of $S - N$ in common,* (2) *N contains r points (r finite) whose sum intersects every element of G,* (3) *$G^* + N$ is closed, and* (4) *the collection G is upper semicontinuous. If $G^* + N$ separates a point A from a point B in S, there exists a subcollection H of G, containing not more than $2r$ elements, such that $H^* + N$ separates A from B in S.*

Suppose first that $G^* + N$ is compact. Let $\{\Delta_i\}$ be a sequence of compact domains closing down on N such that $\bar{\Delta}_i$ $(i = 1, 2, \cdots)$ contains neither A nor B and has only a finite number of components. Consider the collection G_i $(i = 1, 2, \cdots)$ of all elements each of which is the sum of $\bar{\Delta}_i$ and an element of G. By Theorem 7 there exists a finite number of elements of G_i whose sum separates A from B in S. Therefore, by a theorem [9] of the writer, there exist $2r$ or less elements of G_i whose sum separates A from B in S. Denote these elements by $g_i^1, \cdots, g_i^{2r}$. There exists a subsequence $\{i_n\}$ $(n = 1, 2, \cdots)$ of the positive integers such that the sequence $\{g_{i_n}^j\}$ $(j = 1, \cdots, 2r)$ has a sequential limiting set R^j. With the aid of condition (4) of the hypothesis it follows that R^j is a subset of $N + v^j$, where v^j denotes some element of G. Hence $R = (R^1 + \cdots + R^{2r})$ is a subset of $N + (v^1 + \cdots + v^{2r})$. But R separates A from B in S. Hence for H we may take the collection $v^1, \cdots, v^{2r}$.

If $G^* + N$ is not compact, the proof can be made by an inversion of the plane.

Theorem 9 is not true if condition (4) is omitted.

THE UNIVERSITY OF TEXAS.

---

[7] "Théorème sur trois continus," *Monatshefte für Mathematik und Physik*, vol. 36 (1929), pp. 77-80.

[8] "Concerning two internal properties of plane continua," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 670-674.

[9] *Loc. cit.*

THE ANALYSIS OF CERTAIN CURVES BY MEANS OF DERIVED LOCAL SEPARATING POINTS.

By G. E. SCHWEIGERT.

1. *Introduction.* It is well known [1] that the removal of the local separating points [2] from a regular curve leaves a set of dimension zero. This suggests that a similar result may be true for an hereditarily locally connected continuum, that is, a continuum having the property that every subcontinuum is locally connected. However an example by Gehman [3] shows that non-degenerate components may remain in such a continuum after the removal of its local separating points. In this example a line segment is the only component that remains and if we remove in addition the local separating points of the line segment, the result is a set of dimension zero. Thus although it is clearly necessary to remove points other than the local separating points, the situation in this example suggests that *given an hereditarily locally connected continuum, a set of dimension zero may be reached by a countable iteration of the process of removing the local separating points from the components that remain at each successsive stage.* It is to this problem that sections 2, 3, 4, and 5 of this paper are devoted.

2. *Derived local separating points.* For any subset M of a compact, metric, hereditarily locally connected continuum H, let $L_0(M)$ denote the set of all local separating points of M. Then we may write $H - L_0(H) = \Sigma C_k^1 + P_1$ where the C_k^1 are the non-degenerate components in the set on the right and P_1 is a set of dimension zero. In doing this we make use of the fact [4] that an arbitrary subset of an hereditarily locally connected continuum is made up of non-degenerate components and a zero-dimensional set, with the former in the form of a (null) sequence. Now let $L_1(H) = \Sigma L_0(C_k^1)$, and proceed to represent the set $H - [L_0(H) + L_1(H)]$ in a similar fashion. This process

[1] G. T. Whyburn, "Über die Strukture regulärer Kurven," *Wiener Akademie Anzeiger* (1930), Nr. 6.

[2] If p is a cut-point of a connected neighborhood R in a connected and locally connected space then p is said to be a *local separating point*, that is, $R - p = R_1 + R_2$ where $\bar{R}_1 R_2 = 0 = R_1 \bar{R}_2$ and neither R_1 nor R_2 is vacuous.

[3] *Annals of Mathematics*, vol. 27 (1926), p. 43.

[4] G. T. Whyburn, "Concerning hereditarily locally connected continua," *American Journal of Mathematics*, vol. 53, no. 2 (1931), Corollary c, p. 379 and Theorem 4, p. 377.

continues under an inductive scheme of definition which uses the transfinite ordinal numbers. For this purpose we assume that the components C_N^γ have been defined as the components of the set $H - \sum_{\beta<\gamma} L_\beta(H)$ for every $\gamma < \alpha$, and also that $L_\gamma(H) = \Sigma L_0(C_N^\gamma)$ for every $\gamma < \alpha$. Then the general terms are $H - \sum_{\beta<\alpha} L_\beta(H) = \Sigma C_i^\alpha + P_\alpha$ and $L_\alpha(H) = \Sigma L_0(C_i^\alpha)$ where: (a) for each pair (i,α), $i = 1, 2, 3, \cdots$ and α transfinite, C_i^α is a component of $H - \sum_{\beta<\alpha} L_\beta(H)$; (b) P_α is a set of dimension zero.

For each i the set C_i^α is said to be a *derived component of index* α relative to H. Similarly the points in the set $L_\alpha(H)$ are the *derived local separating points of index* α relative to the continuum H. Under this notation H is denoted by C_1^0. Due to a theorem by Wilder [5] it is known that each derived component is a locally connected set.

If we make use of the fact that the derived local separating points are dense in each derived component [6] we may put the solution to our first problem in the form of the next proposition.

THEOREM. *If H is any compact, metric, hereditarily locally connected continuum there exists a number α of the first or second number class such that $L_\alpha(H) = 0$.*

If $\alpha(H)$ is the first such ordinal number α, then $\alpha(H)$ is said to be the *index* of the continuum H.

The above theorem [7] will be established with the aid of the work of Menger and Reschovsky on rational curves.

3. *Unilateral and bilateral limit points.* A generic member of the derived components of index α will be denoted by C^α. Whenever C^α, C^β, C^γ, etc., are used they are to be thought of as members of a monotone sequence which terminates with the component of highest index.

The symbol R denotes any fixed neighborhood in H having a countable boundary and of such a nature that it is in diameter less than that of any component with which it appears. The boundary of any such neighborhood is

---

[5] R. L. Wilder, *Proceedings of the National Academy of Sciences*, vol. 15 (1929), p. 616.

[6] This result is an easy consequence of the fact that any isolated point of an irreducible cutting between two points of a connected and locally connected set M is a local separating point of M. For the theorem quoted here see G. T. Whyburn, *Monatshefte für Mathematik und Physik*, vol. 36 (1929).

[7] A more elementary proof using only the material now at hand together with the fact that H is a rational curve could be given at this point.

a countable cutting of H for which we consider only the particular separation $H - F(R) = R + (H - \bar{R})$. Similarly for the cutting $F(R)C^\alpha$ of C^α we use but one separation $C^\alpha - F(R)C^\alpha = C^\alpha R + C^\alpha(H - \bar{R})$.

If X is a cutting of a set M and the point x belongs to X, then x is said to be a *unilateral limit point of M with respect to the fixed separation* $M - X = M_1 + M_2$ provided x is a limit point of M_1 or M_2 but not of both these sets. If x is a limit point of both M_1 and M_2 then x is a *bilateral limit point of M with respect to the separation.* Since we are to be concerned only with cases in which there is a known separation (as in the paragraph above) the shorter terms unilateral and bilateral limit point *of* M will be used.

It can easily be shown that if x is a unilateral limit point of C^α then x is a unilateral limit point of C^β provided that $x \in F(R)C^\beta$ and $\beta > \alpha$. In regard to the bilateral limit points, what amounts to the converse is true, namely, if x is a bilateral limit point of C^α then x is a bilateral limit point of any component of smaller index.

The α-th derivative of the boundary of any neighborhood N will be denoted by $F(N)^\alpha$. A set E belonging to the cutting $F(R)C^\alpha$ of C^α will be said to be *of type* B^α provided there exists for each point $y \in E$ a number $\alpha(y) < \alpha$ such that y is a unilateral limit point of $C^{\alpha(y)}$. It follows that y is a unilateral limit point of all those derived components of index greater than $\alpha(y)$, including C^α. Another most useful result in this connection is that y is a unilateral limit point of $C^{\alpha-1}$ when α is an isolated number. If E is a set of type B^α we shall write $E = B^\alpha$. The only sets E to be used here are the particular type shown in the next paragraph.

We now prove a lemma from which the principal proposition of this paper is to be deduced. The lemma is stated in terms of the symbols just developed and its second part is a direct consequence of the first.

LEMMA. (1) $[F(R) - F(R)^\zeta]C^\zeta = B^\zeta$; (2) *If x is a bilateral limit point of C^ζ and $x \in [F(R) - F(R)^{\zeta+1}]C^\zeta$, then x is a derived local separating point of index ζ.*

Proof. When $\zeta = 1$ it must be shown in part 1 that there is *no* bilateral limit point x of C^0 (that is, of H) among those isolated points of $F(R)$ which are in C^1. In this connection we recall that the set $F(R)$ is a *countable* cutting of H. This allows us to obtain a separation of H by means of its bilateral limit points alone and from this stage it is but an easy step to show that x is a local separating point of H. Thus x is not a point of C^1 which is contrary to the hypothesis for the existence of that point.

The point x described in part 2 must lie in $F(R)^1 - F(R)^2$ because

it is a bilateral limit point of C^1 and therefore cannot be an isolated point in $F(R)$. But this means that x is an isolated point of X_0, where X_0 is the set of all bilateral limit points of C^1. Then by means of the same argument as was used above it follows that x is a point of $L_1(H)$.

Now for the purpose of induction we assume that the Lemma is true for $\zeta < \eta$ and show that it holds for $\zeta = \eta$.

If we assume the contrary of what is to be shown, we have that some point x is a bilateral limit point of C^ζ for all $\zeta < \eta$ where $x \in [F(R) - F(R)^\eta]C^\eta$. It follows that if η is a limit number, there exists a number $\lambda < \eta$ such that x non-$\in F(R)^\lambda$. But then $x \in [F(R) - F(R)^\lambda]C^\lambda$, and furthermore x is a bilateral limit point of C^λ because that property holds for C^ζ where $\zeta < \eta$. This is impossible in view of the fact that $[F(R) - F(R)^\lambda]C^\lambda = B^\lambda$ holds by assumption and hence no point of the set on the left is a bilateral limit point of C^λ. If on the other hand η is an isolated number, we know that x is a bilateral limit point of $C^{\eta-1}$. Moreover $x \in [F(R) - F(R)^\eta]C^{\eta-1}$, since $x \in C^\eta$. Then by part 2 for $\zeta = \eta - 1$, we find that $x \in L_{\eta-1}(H)$. This is impossible when $x \in C^\eta$.

It remains to be shown that if x is a bilateral limit point of C^η and $x \in [F(R) - F(R)^{\eta+1}]C^\eta$, then $x \in L_\eta(H)$. Certainly such a point x cannot belong to $F(R) - F(R)^\eta$ because of the results just above; hence $x \in F(R)^\eta - F(R)^{\eta+1}$. Now let X_0 be the set of all bilateral limit points of C^η. Then $X_0[F(R) - F(R)^\eta]C^\eta = 0$ because the last two factors form a set of type B^η. It follows that x is an isolated point of X_0 and hence by the argument used earlier that $x \in L_\eta(H)$. This completes the induction, that is, the proof of the Lemma.

4. *Relation between index and genus.* A point p of a set M is said to be of *genus* α *in* M [8] provided: (a) there exists arbitrarily small neighborhoods in M closing down on p in such a way that the α-th derivative of the boundary of each neighborhood is vacuous; (b) for some $\epsilon > 0$ and every $\beta < \alpha$ there exists *no* neighborhood in M which contains p, is of diameter less than ϵ and has the further property that the β-th derivative of the boundary is vacuous. The genus of the point p will be denoted by $g(p)$.

Suppose there is a point p in a derived component C^α such that $g(p) \leq \alpha$. Then since p is also in $C^{g(p)}$, there exists a neighborhood R in H such that R contains p and does not contain all points of $C^{g(p)}$. Moreover R can be so chosen that $F(R)^{g(p)} = 0$ and $F(R)$ is countable. It follows from the fact that $C^{g(p)}$ is connected that there exists a point x of $F(R)C^{g(p)}$ such that x is a bilateral limit point of $C^{g(p)}$. Thus $x \in [F(R) - F(R)^{g(p)}]C^{g(p)}$ because

[8] *Geschlecht*—definition due to Menger. See *Kurventheorie* (Teubner), p. 294.

$F(R)^{g(p)} = 0$. This is contrary to the Lemma in section 3, hence we have shown that the following proposition is true.

THEOREM. *If H is any compact, metric, hereditarily locally connected continuum and C^α is any one of its derived components of index α, then each point of C^α is of genus greater than α.*

5. *Solution to the problem stated in section* 1. It has been shown by Reschovsky [9] that there is associated with each rational curve K a number α of the first or second class such that the genus of each point of K is less than or equal to α. The smallest number $g(K)$ for which this statement holds is called the genus of the rational curve.[10]

This classification applies to the hereditarily locally connected continuum H for it has been established by G. T. Whyburn [11] that each such continuum is a rational curve. Thus it is an obvious corollary to the Theorem immediately above that $g(H) \geq \alpha(H)$. This means that $\alpha(H)$ is also of the first or second number class, consequently the Theorem stated in section 2 is proven and thus we have a positive solution to our original problem.

6. *Regular and rational bases.* A subset B of an arbitrary set M is said to be a *regular basis* for M provided that, (1) each regular point of M is contained in arbitrarily small neighborhoods N_ϵ in M which close down on p in such a way that $F(N_\epsilon)$ is finite and is contained in B, (2) when p is of order k, the neighborhoods N_ϵ can be chosen so that $F(N_\epsilon)$ consists of exactly k points of B. Similarly B is said to be a *rational basis* for M provided that each rational point of M is contained in arbitrarily small neighborhoods in M which close down on that point and have countable boundaries in B. Obviously a set of dimension zero remains on removing a rational (regular) basis from a rational (regular) curve.

In the paper by G. T. Whyburn referred to in the introduction it is shown that B will be a countable regular basis for a regular curve K provided that $B = T + D$, where T is the set of all local separating points of order greater than 2 in K and D is a countable set which is dense on each free arc in K. It is our purpose to establish a proposition which is analogous to this one in that it enables us to choose a rational basis for a curve H from its derived local separating points. The next Theorem is offered as a solution to this problem.

THEOREM. *Let H be any compact, metric, hereditarily locally connected*

[9] Reschovsky, *Fundamenta Mathematica*, vol. 15 (1930), p. 18.

[10] *Loc. cit.*

[11] "Concerning hereditarily locally connected continua."

continuum and let B be any subset of H such that $B = \sum_{i,a} T_i^a + D$, where D is a countable set, and such that in regard to each derived component C_i^a: (a) T_i^a is the set of all derived local separating points of order greater than two in that component and of index equal to that of the component, (b) *D is dense on each free arc. Then B is a countable rational basis for H and the part of B which is contained in each derived component is a countable regular basis for that component.*

In the statement of this Theorem we again follow the convention that H is uniquely the derived component of index zero; hence when H is a regular curve, this Theorem reduces to the one quoted in the above paragraph. The proof of this Theorem is an almost direct result of the following Lemma.

LEMMA. *If C is any connected and locally connected set and B is any subset of C such that $B = T + D$, where T is the set of all local sepaarting points of order greater than two in C and D is a countable set which is dense on each free arc in C, then B is a regular basis for C.* Moreover B is a *countable* regular basis provided the nodules in C form a null sequence.

This Lemma follows from an argument based on the work of G. T. Whyburn.[12]

Since each derived component C_i^a of the curve H is locally connected we may apply the Lemma to each of these components and obtain a regular basis $B_i^a = T_i^a + D_i^a$. This basis is a countable set in each case because the nodules are certainly in a null sequence when the containing space is hereditarily locally connected. Let T and D denote the sets $\sum_{i,a} T_i^a$ and $\sum_{i,a} D_i^a$ respectively and denote their sum by B. From this approach it is obvious that the part of B that is contained in each derived component is a regular basis for that component. Moreover there are but a countable number of derived components hence B is a countable set. Thus in order to complete the proof of the Theorem we have to show that B is a rational basis.

It is first shown that the non-degenerate components of $C_i^a - B_i^a$ are exactly the components of $C_i^a - L_0(C_i^a) = \Sigma C_k^{a+1} + P_{a+1}$ after which it is immediate that $H - B$ is a set of dimension zero. For the purpose of comparing the two sets, let $C_i^a - B_i^a = \Sigma \Gamma_j^{a+1} + Q_{a+1}$ where the Γ_j^{a+1} are the non-degenerate components of the containing set and Q_{a+1} is a set of dimension zero. Now assume there exists a point p in some fixed component Γ_j^{a+1} such that p non-$\epsilon\, C_k^{a+1}$ for all k. Then p is either a point of $L_0(C_i^a) - B_i^a$ or a point of P_{a+1}. When p is in the former set, let $q \epsilon \Gamma_j^{a+1}$ be a point distinct

[12] See his papers "On the structure of connected and connected im kleinem point sets," *Transactions of the American Mathematical Society*, vol. 32 (1930), pp. 926-943, and "Über die Struckture regulärer Kurven," *Wiener Anzeiger* (1930), Nr. 6.

from p. Then there exists a neighborhood R in C_i^a such that R contains p but does not contain q and such that $F(R) = x, y \epsilon B_i^a$. (The point p is of order two by hypothesis.) However Γ_j^{a+1} is connected and consequently $F(R)$ contains a point z which is distinct from x and y. This is a contradiction. When p is in the latter set P_{a+1}, the point p is contained in arbitrarily smali neighborhoods in $C_i^a - L_0(C_i^a)$ each of which does not contain all of that set and has a vacuous boundary. On the other hand we have just shown that Γ_j^{a+1} can contain no point of $L_0(C_i^a) - B_i^a$ hence $\Gamma_j^{a+1} - p$ is contained in a component C_k^{a+1} for some k. These two situations are inconsistent and bring us to the result that $H - B$ is a set of dimension zero. But it obviously follows that $(H - B) + p$, where p is any point of B, is also a set of dimension zero and consequently B is a rational basis. This completes the proof of the Theorem.

It is to be noted that in each case where a basis is mentioned in this section, that basis is prescribed, except for a certain freedom in the choice of the sets which are dense on the free arcs, by the set which contains it. Thus it may be said that each basis is a *natural* basis.

7. *Conclusion.* We are now in a position to demonstrate that any hereditarily locally connected continuum H is topologically contained in one which is of index 1. For this purpose we need the fact that there exists a locally connected Universal Curve in Three-space which contains a subset topologically equivalent to each compact, metric curve.[13] Now let H be any hereditarily locally connected continuum and let T be a homeomorphism throwing H into a subset $T(H)$ of the Universal Curve. Let $T(B) = b_1 + b_2 + b_3 + \cdots$ be the image under T of a countable rational basis B in H. Then add to $T(H)$ a line segment $L(1)$ of a length 1 having the point b_1 in common with $T(H)$. In general $L(n)$ is to be a member of a monotone sequence of disjoint line segments which meets $T(H)$ in just one point $b_n \epsilon T(B)$ and has the length $1/n$. The curve $T(H) + \Sigma L(n)$ will then be of index 1.

One can easily see that if the original curve H is contained in a space of a finite number of dimensions, it can be augmented in the manner described above and the resulting curve will have the same genus as the original curve. This suggests to the author that, while some examples of his (similar to the one by Gehman) showing plane curves with indices $2, 3, 4, \cdots, \omega$ are cases for which $\alpha(H) = g(H)$, this method will alter those examples to show that there are cases in which $\alpha(H)$ takes the full range allowed by the inequality $\alpha(H) \leq g(H)$. In contrast to this we know that if H is a regular curve $\alpha(H) = g(H) = 1$.

THE JOHNS HOPKINS UNIVERSITY.

[13] Menger, *Kurventheorie*, p. 365.

ON THE INDEPENDENCE OF HILBERT AND ACKERMANN'S POSTULATES FOR THE CALCULUS OF PROPOSITIONAL FUNCTIONS.

J. C. C. McKINSEY.

Hilbert and Ackermann have given a set of postulates for the calculus of propositional functions,[1] but they state[2] that the independence of the set has not been investigated. In this paper I show that the postulates in question are independent.

I first set forth the undefined marks and the postulates of the system. The undefined marks[3] are:

1) Propositional-variables: $X, Y, Z, \cdots$.

2) Individual-variables: $x, y, z, \cdots$.

3) Function-variables: $F(\ ), G(\ ), \cdots$.

4) The signs (x), (y), etc. These can occur before a function-variable into which has been substituted an individual-variable: thus, $(x)F(x)$, etc.

5) The signs (Ex), (Ey), etc. Giving $(Ex)F(x)$, etc.

6) The sign $\vee$.

7) The sign $-$.

In the expression $(\alpha)\beta$ the sign (α) is called the *universal-operator*; and in $(E\alpha)\beta$, the sign $(E\alpha)$ is called the *existential-operator*. Universal-operators and existential-operators are together referred to as *operators*. In both operators the α is always an individual-variable, and it is termed the *operator-variable*. β is called the *scope* of α. If an individual-variable x occurs within the scope of an operator having x as operator-variable, then x is called an *apparent* variable, or a *bound* variable. Other individual-variables are called *real* variables, or *free* variables.

[1] D. Hilbert and W. Ackermann, *Grundzüge der Theoretischen Logik*, Julius Springer, 1928, pp. 53-54.

[2] *Loc. cit.*, p. 68. Recently, however, a set of independence examples for a closely related set of postulates has been given by B. Notcutt, "A set of independent postulates for propositional functions of one variable," *Annals of Mathematics*, vol. 37 (1935), pp. 670-678. The similarity of Notcutt's postulates to Hilbert and Ackermann's arises from the fact that both sets are modifications of the set given in *Principia Mathematica*.

[3] The undefined terms are not given by Hilbert and Ackermann in the usual mathematical form; that is to say, the authors speak simply of undefined "marks," instead of undefined "classes," "operations," and "relations."

The term "expression" (as used in the postulates) is precisely defined as follows:

1. A sign for a proposition (as X, Y, etc.) is an expression.

2. A sign for a propositional function in which each blank space is filled either by an individual-variable, or by the proper name of an individual is an expression.

3. If A is an expression, so is $\bar{A}$; and, in case the free variable x occurs in A, so is $(x)A$ and $(Ex)A$.

4. If A and B are expressions, so is $A \vee B$.

5. But the restriction is to be understood that in an expression we never have overlapping scopes of a like-indicated variable. Thus, for example, $(x)F(x) \vee (Ex)G(x)$ is an expression, but $(x)(y)[F(x) \vee (Ex)G(x, y)]$ is not.

The complex sign $A \rightarrow B$ is defined to be $\bar{A} \vee B$. From 3 and 4, it is seen that $A \rightarrow B$ is an expression if A and B are expressions.

On these undefined marks the following postulates[4] are imposed:

a) $X \vee X \rightarrow X$.
b) $X \rightarrow X \vee Y$.
c) $X \vee Y \rightarrow Y \vee X$.
d) $(X \rightarrow Y) \rightarrow [Z \vee X \rightarrow Z \vee Y]$.
e) $(x)F(x) \rightarrow F(y)$.
f) $F(y) \rightarrow (Ex)F(x)$.

α) (*Rule of Substitution*). If in any correct formula we substitute any expression for a propositional-variable (making the same substitution, however, wherever this one propositional-variable occurs) then the resulting formula is also correct; similarly, for a function-variable with arguments $x, y, \cdots, u$ we may substitute an expression which depends on $x, y, \cdots, u$ (or, if desired, on still more variables). An individual-variable may be replaced by an individual-variable denoted by a different letter, or by a proper name from the range of the variable.[5]

[4] Postulate a) may be understood as reading "$X \vee X \rightarrow X$ is a correct formula"; and similarly for b)–f).

[5] Postulate α) is followed by the remark that it is to be understood that, when a substitution is made into an expression within the scope of an apparent variable, the expression substituted should not contain the apparent variable in question. Thus, for example, we cannot go from the correct formula $Z \vee (x)F(x) \rightarrow (x)[Z \vee F(x)]$ to the formula $G(x) \vee (x)F(x) \rightarrow (x)[G(x) \vee F(x)]$.

β) (*Rule of Inference*). If A is a correct formula and $A \rightarrow B$ is a correct formula, then B is a correct formula.

γ) Let $B(x)$ be any expression which depends on x, and let A be any expression which does not depend on x. Then if $A \rightarrow B(x)$ is a correct formula, $A \rightarrow (x)B(x)$ is also a correct formula. And if $B(x) \rightarrow A$ is a correct formula, $(Ex)B(x) \rightarrow A$ is also a correct formula.

I here add certain lemmas which will be used in the independence proofs.

LEMMA I. *The following formula is deducible from Hilbert's postulates:*

$$F(y) \rightarrow F(y).$$

Proof.

(1)	$\left[\text{d}), \frac{X \vee X}{X}, \frac{X}{Y}, \frac{\bar{X}}{Z}\right]$	$(X \vee X \rightarrow X) \rightarrow [\bar{X} \vee (X \vee X) \rightarrow (\bar{X} \vee X)]$
(2)	[a)]	$X \vee X \rightarrow X$
(3)	[(1), (2), β)]	$\bar{X} \vee (X \vee X) \rightarrow (\bar{X} \vee X)$
(4)	[(3), Def. of $\rightarrow$]	$(X \rightarrow X \vee X) \rightarrow (X \rightarrow X)$
(5)	$\left[\text{b}), \frac{X}{Y}\right]$	$X \rightarrow X \vee X$
(6)	[(4), (5), β)]	$X \rightarrow X$
	$\left[(6), \frac{F(y)}{X}\right]$	$F(y) \rightarrow F(y).$

LEMMA IIa. *If K_1 is independent of x, and $K_2(x)$ depends on x, and if $K_1 \rightarrow K_2(x)$ is deducible from Hilbert's postulates, then $K_1 \rightarrow (Ex)K_2(x)$ is also deducible from Hilbert's postulates.*

LEMMA IIb. *Similarly, if $K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates, then $(x)K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates.*

Proof. I first show that if $A \rightarrow B$ and $B \rightarrow C$ are deducible from Hilbert's postulates, where A, B, and C are any expressions, then $A \rightarrow C$ is deducible also. For substituting B for X, C for Y, and $\bar{A}$ for Z in postulate d), and applying the definition of $\rightarrow$, we get

$$(B \rightarrow C) \rightarrow [(A \rightarrow B) \rightarrow (A \rightarrow C)].$$

Then since $B \rightarrow C$ is a correct formula, we have, by postulate β)

$$(A \rightarrow B) \rightarrow (A \rightarrow C).$$

Then since $A \rightarrow B$ is a correct formula, we have, again by postulate β), that $A \rightarrow C$ is a correct formula.

Now suppose we have $K_1 \rightarrow K_2(x)$ given as being deducible from Hilbert's postulates. Substituting $K_2(x)$ for $F(x)$ in postulate f), we have

$K_2(x) \rightarrow (Ex)K_2(x)$. Hence by the first part of the proof, the formula $K_1 \rightarrow (Ex)K_2(x)$ is deducible from Hilbert's postulates.

Similarly, if we have $K_2(x) \rightarrow K_1$, then, since by postulate e) we also have $(x)K_2(x) \rightarrow K_2(x)$, it follows that $(x)K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates.

LEMMA III. *The following formulas are not deducible from Hilbert's postulates:*

a) $(Ex)F(x) \rightarrow F(y)$
b) $F(y) \rightarrow (x)F(x)$.

Proof. The following example is given by Hilbert in another connection.[6] A formula is said to be satisfied if by the following process it is transformed into a correct formula of the calculus of propositions: First generalize all the free variables by putting the corresponding universal-operators at the left of the formula. Then[7] if $A(x)$ is any expressions, substitute $A(0)$ & $A(1)$ for $(x)A(x)$, and substitute $A(0) \vee A(1)$ for $(Ex)A(x)$. Continue this process until all apparent variables have been eliminated. The formula will then contain only propositional variables and symbols of the form $F(0)$ and $F(1)$. Then substitute for $F(0)$, wherever it occurs, the same propositional variable, for $F(1)$ another, and so on. By this process, postulate e), for example, becomes successively:

$$(y)[(x)F(x) \rightarrow F(y)]$$
$$[(x)F(x) \rightarrow F(0)] \,\&\, [(x)F(x) \rightarrow F(1)]$$
$$[F(0) \,\&\, F(1) \rightarrow F(0)] \,\&\, [F(0) \,\&\, F(1) \rightarrow F(1)]$$
$$[X \,\&\, Y \rightarrow X] \,\&\, [X \,\&\, Y \rightarrow Y].$$

The last formula is a correct formula of the calculus of propositions. Thus postulate e) holds for this interpretation. In a similar way it may be shown that all the other postulates are satisfied by this interpretation.

Consider now the formula $(Ex)F(x) \rightarrow F(y)$. This becomes successively:

$$(y)[(Ex)F(x) \rightarrow F(y)]$$
$$[(Ex)F(x) \rightarrow F(0)] \,\&\, [(Ex)F(x) \rightarrow F(1)]$$
$$[F(0) \vee F(1) \rightarrow F(0)] \,\&\, [F(0) \vee F(1) \rightarrow F(1)]$$
$$[X \vee Y \rightarrow X] \,\&\, [X \vee Y \rightarrow Y].$$

That this is not a correct formula of the calculus of propositions may be seen by taking X and Y of opposite truth-values. Hence the formula $(Ex)F(x) \rightarrow F(y)$ is not deducible from Hilbert's postulates.

[6] *Loc. cit.*, p. 66.

[7] $A \,\&\, B$ is defined as $[\bar{A} \vee \bar{B}]$.

Similarly, the formula $F(y) \rightarrow (x)F(x)$ becomes successively:

$$(y)[F(y) \rightarrow (x)F(x)]$$
$$[F(0) \rightarrow (x)F(x)] \,\&\, [F(1) \rightarrow (x)F(x)]$$
$$[F(0) \rightarrow F(0) \,\&\, F(1)] \,\&\, [F(1) \rightarrow F(0) \,\&\, F(1)]$$
$$[X \rightarrow X \,\&\, Y] \,\&\, [Y \rightarrow X \,\&\, Y]$$

and this also fails when X and Y have opposite truth-values.

I now proceed to prove that Hilbert's postulates are independent. The proofs are given, not in the exact order in which the postulates are listed, but rather fall into two groups, depending on the nature of the proof-methods used. The first group comprises examples for postulates a), b), c), d), β), and α); and the second group for postulates e), f), and γ).

For all the independence proofs of the first group, I interpret $(E\alpha)\beta$ in a formula as meaning β, and also $(\alpha)\beta$ as meaning β; so that all apparent variables are eliminated from the formula. I then interpret all real variables as meaning a certain fixed individual, which I denote by the number 7; so that the formula now contains only propositional-variables, and function-variables with the argument 7; e. g. $F(7)$, $G(7)$, etc. I then interpret $F(7)$ as meaning a propositional-variable. Thus all formulas reduce to formulas in terms of the undefined ideas of the calculus of propositions. For example, the formula

$$(x)F(x) \vee (y)G(y) \rightarrow F(z) \vee (Ew)G(w)$$

becomes first

$$F(x) \vee G(y) \rightarrow F(z) \vee G(w)$$

and then

$$F(7) \vee G(7) \rightarrow F(7) \vee G(7)$$

and finally

$$X \vee Y \rightarrow X \vee Y.$$

The operations $X \vee Y$ and $\bar{X}$ (and hence $X \rightarrow Y$) are then defined by tables, in the usual way. And I say a formula is satisfied by the interpretation if it is such as to give the value 0 for all possible values of the propositional-variables.

Tables[8] *for Postulate* a)

$\vee$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	0

X	$\bar{X}$
0	1
1	0
2	2

$\rightarrow$	0	1	2
0	0	1	2
1	0	0	0
2	0	2	0

[8] The tables used for postulates a), c), and d) are given by Hilbert and Ackermann in another connection. *Loc. cit.*, pp. 31-33.

(Here propositional-variables can take on the values 0, 1, 2.)

Postulate a) fails when $X = 2$. For $2 \vee 2 \rightarrow 2$ becomes $0 \rightarrow 2$, which is 2.

Postulates b), c), and d) are satisfied.

Postulate e) becomes, successively:

$$F(x) \rightarrow F(y)$$
$$F(7) \rightarrow F(7)$$
$$X \rightarrow X,$$

and we see from the table for $X \rightarrow Y$ that $X \rightarrow X$ always has the value 0.

Postulate f), similarly, becomes $X \rightarrow X$.

Postulate α) is satisfied since any expression reduces to a complex involving only propositional variables and the signs $\vee$ and $-$; and, since the tables define class-closing operations, such a complex cannot take on *more* (though, of course, it may take on fewer) values than the propositional-variable for which it is substituted.

To see that postulate β) is satisfied, we need only observe that the table for $X \rightarrow Y$ never gives the value 0 when $X = 0$ and $Y \neq 0$.

Postulate γ) is satisfied since $A \rightarrow B(x)$ and $A \rightarrow (x)B(x)$ reduce to the same thing; and similarly for $B(x) \rightarrow A$ and $(Ex)B(x) \rightarrow A$.

Thus all the postulates except postulate a) are satisfied, so a) is independent of the other postulates.

Tables[9] *for Postulate* b)

$\vee$	0	1	2	3
0	0	0	0	0
1	0	3	0	3
2	0	0	2	2
3	0	3	2	3

X	$\bar{X}$
0	3
1	2
2	1
3	0

$\rightarrow$	0	1	2	3
0	0	3	2	3
1	0	0	2	2
2	0	3	0	3
3	0	0	0	0

(Propositional-variables take on values 0, 1, 2, 3.)

Postulate b) fails for $X = Y = 1$; since $1 \rightarrow 1 \vee 1$ becomes $1 \rightarrow 3$, which is 2.

The other postulates can be shown to hold as in the discussion for postulate a).

---

[9] The tables for postulate b) are isomorphic to certain tables given by Wajsberg to show the consistency of C. I. Lewis' system of "strict implication" (see Lewis and Langford, *Symbolic Logic*, p. 493, Group II), where $\bar{X}$ corresponds to $\sim p$, $X \rightarrow Y$ to $p \rightarrow q$, and $X \vee Y$ to $\sim p \rightarrow q$. Postulate b) yields one of the so-called "paradoxes of material implication."

Tables for Postulate c)

$\vee$	0	1	2
0	0	0	0
1	0	1	1
2	0	0	2

X	$\bar{X}$
0	1
1	2
2	0

$\rightarrow$	0	1	2
0	0	1	1
1	0	0	2
2	0	0	0

(Propositional-variables take on values 0, 1, 2.)

Postulate c) fails for $X = 2$, $Y = 1$.

Tables for Postulate d)

$\vee$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	2	0
3	0	3	0	3

X	$\bar{X}$
0	1
1	0
2	3
3	0

$\rightarrow$	0	1	2	3
0	0	1	2	3
1	0	0	0	0
2	0	3	0	3
3	0	0	0	0

(Propositional-variables take on values 0, 1, 2, 3.)

Postulate d) fails for $X = 3$, $Y = 1$, $Z = 2$.

Tables for Postulate β)

$\vee$	0	1
0	0	0
1	0	1

X	$\bar{X}$
0	0
1	0

$\rightarrow$	0	1
0	0	0
1	0	0

(Propositional variables take on the values 0, 1.)

Postulate β) fails. For the formula,

$$(X \rightarrow Y) \rightarrow (X \vee Y)$$

gives 0 for all values of X, Y; and the formula $X \rightarrow Y$ gives 0 for all values of X, Y; but the formula $X \vee Y$ gives the value 1 for $X = Y = 1$.

The other postulates are easily seen to hold.

Tables for Postulate α)

$\vee$	0	1	2
0	0	0	1
1	0	1	1
2	0	1	1

X	$\bar{X}$
0	2
1	0
2	0

$\rightarrow$	0	1	2
0	0	1	1
1	0	0	1
2	0	0	1

(Propositional-variables take on the values 0, 1. Propositional-variables do *not* take on the value 2.)

Postulate α) fails. For it is easily seen that the formula $X \vee X \rightarrow X$ is satisfied for $X = 0, 1$; and if postulate α) held, then the formula $\bar{X} \vee \bar{X} \rightarrow \bar{X}$ (got by substituting $\bar{X}$ for X in $X \vee X \rightarrow X$) should also be satisfied. But $\bar{X} \vee \bar{X} \rightarrow \bar{X}$ is not satisfied; for if we take $X = 0$, it becomes $\bar{0} \vee \bar{0} \rightarrow \bar{0}$, which is $2 \vee 2 \rightarrow 2$, which is $1 \rightarrow 2$, which is 1. Hence α) is not satisfied.

The other postulates are readily shown to hold (remembering that the propositional-variables do not assume the value 2).

I come now to the independence-proofs of the second group. The general method used here may be described as follows. Suppose we had a set of postulates P_1, P_2 and another set Q_1, Q_2. If T represents a transformation of the undefined ideas of P_1, P_2, I designate the respective transforms of P_1 and P_2 by $T(P_1)$ and $T(P_2)$. Suppose we are able to find a transformation T_1 such that $T_1(P_1)$ is deducible from Q_1, Q_2, while $T_1(P_2)$ is independent of Q_1, Q_2; then P_2 is independent of P_1. For if P_2 were deducible from P_1, then $T_1(P_2)$ would be deducible from $T_1(P_1)$, and hence from Q_1, Q_2, contrary to hypothesis. This proof still holds, moreover, even if we suppose P_1 identical with Q_1 and P_2 identical with Q_2.

Thus, to prove the independence of postulate e), we describe a certain transformation T_e to be applied to the undefined marks of the formulas of Hilbert's system. If $T_e(A)$ is deducible from Hilbert's postulates, we say that A is satisfied; if $T_e(A)$ is independent of Hilbert's postulates, we say that A is not satisfied. (To avoid prolixity, in what follows I shall write "A is H" for "A is deducible from Hilbert's postulates.") The independence-proofs for postulates f) and γ) are similar, except that transformations T_f and T_γ are there used instead of the transformation T_e.

Independence of Postulate e)

Transformation T_e. Replace all universal-operators in the formula by the corresponding existential-operators. Thus, for example, the formula

$$(x)(y)F(x,y) \rightarrow (\exists x)(\exists y)F(x,y)$$

goes over into

$$(\exists x)(\exists y)F(x,y) \rightarrow (\exists x)(\exists y)F(x,y).$$

Postulate e) fails. For, under T_e the formula $(x)F(x) \rightarrow F(y)$ becomes $(Ex)F(x) \rightarrow F(y)$, which, by Lemma IIIa, is not H.

Postulates a), b), c), d), and f) hold, since they are invariant under T_e.

To show that postulate α) holds: Let A be any formula which is satisfied, so that $T_e(A)$ is H. Let E be an expression of the sort specified by α); it is clear that $T_e(E)$ is also an expression of the sort specified by α). Then α) states that if we substitute E in A (obtaining A_1, say) then $T_e(A_1)$ is H. But $T_e(A_1)$ can also be obtained by substituting $T_e(E)$ in $T_e(A)$. Hence since $T_e(A)$ is H, and $T_e(E)$ is an expression of the sort specified in α), it follows that $T_e(A_1)$ is H. Hence A_1 is satisfied. Hence postulate α) holds.

To show that β) holds. Suppose that A and $A \rightarrow B$ are satisfied. Then $T_e(A)$ and $T_e(A \rightarrow B)$ are H. But $T_e(A \rightarrow B)$ is the same as $T_e(A) \rightarrow T_e(B)$. Hence $T_e(B)$ is H, so B is satisfied.

To show that γ) holds. Suppose $A \rightarrow B(x)$ is satisfied. Then $T_e(A \rightarrow B(x))$, and hence $T_e(A) \rightarrow T_e(B(x))$, is H. Hence by Lemma IIa, $T_e(A) \rightarrow (Ex)T_e(B(x))$ is H. But this is the same as $T_e(A) \rightarrow T_e((x)B(x))$, so $T_e(A) \rightarrow T_e((x)B(x))$ is H; so $T_e(A \rightarrow (x)B(x))$ is H, $A \rightarrow (x)B(x)$ is satisfied, and hence the first part of postulate γ) holds. To show that the second part of γ) holds, we suppose that $B(x) \rightarrow A$ is satisfied. Then $T_e(B(x) \rightarrow A)$, and hence $T_e(B(x)) \rightarrow T_e(A)$, is H. So $(Ex)T_e(B(x)) \rightarrow T_e(A)$, and hence $T_e((Ex)B(x) \rightarrow A)$, is H. Hence $(Ex)B(x) \rightarrow A$ is satisfied.

Independence of Postulate f)

Transformation T_f. Replace all existential-operators by the corresponding universal-operators.

Postulate f) fails. For under T_f the formula $F(y) \rightarrow (Ex)F(x)$ becomes $F(y) \rightarrow (x)F(x)$, which, by Lemma IIIb, is not H.

Postulates a), b), c), d), and e) hold, since they are invariant under T_f.

The proof that α) and β) hold is the same as in the example for the independence of e).

The proof that γ) holds is very like the corresponding proof in the discussion for postulate e); Lemma IIb, however, is needed instead of Lemma IIa.

Independence of Postulate γ)

Transformation T_γ. Replace all the free variables in the formula by a single free variable (which is not, however, to be the same as any of the bound variables). Thus instead of $(x)F(x) \rightarrow F(y) \,\&\, F(x)$, we write $(x)F(x) \rightarrow F(y) \,\&\, F(y)$.

Postulate γ) fails. For $F(y) \rightarrow F(x)$ is satisfied, since $T_\gamma(F(y) \rightarrow F(x))$ is $F(y) \rightarrow F(y)$, and this last is H by Lemma I; but $F(y) \rightarrow (x)F(x)$ is not satisfied, since it is invariant under T_γ and, by Lemma IIIb, is not H.

Postulates a), b), c), d), e), and f) are satisfied, since they are invariant under T_γ.

The proof that postulates α) and β) are satisfied is the same as the corresponding proof in the discussion of the independence of postulate e).

This completes the proof of the independence of the nine postulates.[10]

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

---

[10] It is, however, perhaps worth mentioning that the postulates could be reduced in number by introducing one of the following definitions:

$$(Ex)F(x) = \overline{[(x)\overline{F(x)}]} \quad \text{Def.}$$

or

$$(x)F(x) = \overline{[(Ex)\overline{F(x)}]} \quad \text{Def.}$$

AN UNSOLVABLE PROBLEM OF ELEMENTARY NUMBER THEORY.[1]

By ALONZO CHURCH.

1. Introduction. There is a class of problems of elementary number theory which can be stated in the form that it is required to find an effectively calculable function f of n positive integers, such that $f(x_1, x_2, \cdots, x_n) = 2$ [2] is a necessary and sufficient condition for the truth of a certain proposition of elementary number theory involving $x_1, x_2, \cdots, x_n$ as free variables.

An example of such a problem is the problem to find a means of determining of any given positive integer n whether or not there exist positive integers x, y, z, such that $x^n + y^n = z^n$. For this may be interpreted, required to find an effectively calculable function f, such that $f(n)$ is equal to 2 if and only if there exist positive integers x, y, z, such that $x^n + y^n = z^n$. Clearly the condition that the function f be effectively calculable is an essential part of the problem, since without it the problem becomes trivial.

Another example of a problem of this class is, for instance, the problem of topology, to find a complete set of effectively calculable invariants of closed three-dimensional simplicial manifolds under homeomorphisms. This problem can be interpreted as a problem of elementary number theory in view of the fact that topological complexes are representable by matrices of incidence. In fact, as is well known, the property of a set of incidence matrices that it represent a closed three-dimensional manifold, and the property of two sets of incidence matrices that they represent homeomorphic complexes, can both be described in purely number-theoretic terms. If we enumerate, in a straightforward way, the sets of incidence matrices which represent closed three-dimensional manifolds, it will then be immediately provable that the problem under consideration (to find a complete set of effectively calculable invariants of closed three-dimensional manifolds) is equivalent to the problem, to find an effectively calculable function f of positive integers, such that $f(m, n)$ is equal to 2 if and only if the m-th set of incidence matrices and the n-th set of incidence matrices in the enumeration represent homeomorphic complexes.

Other examples will readily occur to the reader.

---

[1] Presented to the American Mathematical Society, April 19, 1935.

[2] The selection of the particular positive integer 2 instead of some other is, of course, accidental and non-essential.

The purpose of the present paper is to propose a definition of effective calculability[8] which is thought to correspond satisfactorily to the somewhat vague intuitive notion in terms of which problems of this class are often stated, and to show, by means of an example, that not every problem of this class is solvable.

2. Conversion and λ-definability. We select a particular list of symbols, consisting of the symbols { , }, (,), λ, [,], and an enumerably infinite set of symbols *a, b, c,* · · · to be called *variables.* And we define the word *formula* to mean any finite sequence of symbols out of this list. The terms *well-formed formula, free variable,* and *bound variable* are then defined by induction as follows. A variable x standing alone is a well-formed formula and the occurrence of x in it is an occurrence of x as a free variable in it; if the formulas F and X are well-formed, $\{F\}(X)$ is well-formed, and an occurrence of x as a free (bound) variable in F or X is an occurrence of x as a free (bound) variable in $\{F\}(X)$; if the formula M is well-formed and contains an occurrence of x as a free variable in M, then $\lambda x[M]$ is well-formed, any occurrence of x in $\lambda x[M]$ is an occurrence of x as a bound variable in $\lambda x[M]$, and an occurrence of a variable y, other than x, as a free (bound) variable in M is an occurrence of y as a free (bound) variable in $\lambda x[M]$.

---

[8] As will appear, this definition of effective calculability can be stated in either of two equivalent forms, (1) that a function of positive integers shall be called effectively calculable if it is λ-definable in the sense of § 2 below, (2) that a function of positive integers shall be called effectively calculable if it is recursive in the sense of § 4 below. The notion of λ-definability is due jointly to the present author and S. C. Kleene, successive steps towards it having been taken by the present author in the *Annals of Mathematics,* vol. 34 (1933), p. 863, and by Kleene in the *American Journal of Mathematics,* vol. 57 (1935), p. 219. The notion of recursiveness in the sense of § 4 below is due jointly to Jacques Herbrand and Kurt Gödel, as is there explained. And the proof of equivalence of the two notions is due chiefly to Kleene, but also partly to the present author and to J. B. Rosser, as explained below. The proposal to identify these notions with the intuitive notion of effective calculability is first made in the present paper (but see the first footnote to § 7 below).

With the aid of the methods of Kleene (*American Journal of Mathematics,* 1935), the considerations of the present paper could, with comparatively slight modification, be carried through entirely in terms of λ-definability, without making use of the notion of recursiveness. On the other hand, since the results of the present paper were obtained, it has been shown by Kleene (see his forthcoming paper, "General recursive functions of natural numbers") that analogous results can be obtained entirely in terms of recursiveness, without making use of λ-definability. The fact, however, that two such widely different and (in the opinion of the author) equally natural definitions of effective calculability turn out to be equivalent adds to the strength of the reasons adduced below for believing that they constitute as general a characterization of this notion as is consistent with the usual intuitive understanding of it.

We shall use heavy type letters to stand for variable or undetermined formulas. And we adopt the convention that, unless otherwise stated, each heavy type letter shall represent a well-formed formula and each set of symbols standing apart which contains a heavy type letter shall represent a well-formed formula.

When writing particular well-formed formulas, we adopt the following abbreviations. A formula $\{\boldsymbol{F}\}(\boldsymbol{X})$ may be abbreviated as $\boldsymbol{F}(\boldsymbol{X})$ in any case where $\boldsymbol{F}$ is or is represented by a single symbol. A formula $\{\{\boldsymbol{F}\}(\boldsymbol{X})\}(\boldsymbol{Y})$ may be abbreviated as $\{\boldsymbol{F}\}(\boldsymbol{X}, \boldsymbol{Y})$, or, if $\boldsymbol{F}$ is or is represented by a single symbol, as $\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{Y})$. And $\{\{\{\boldsymbol{F}\}(\boldsymbol{X})\}(\boldsymbol{Y})\}(\boldsymbol{Z})$ may be abbreviated as $\{\boldsymbol{F}\}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$, or as $\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$, and so on. A formula $\lambda \boldsymbol{x}_1[\lambda \boldsymbol{x}_2[\cdot \cdot \cdot \lambda \boldsymbol{x}_n[\boldsymbol{M}] \cdot \cdot \cdot]]$ may be abbreviated as $\lambda \boldsymbol{x}_1 \boldsymbol{x}_2 \cdot \cdot \cdot \boldsymbol{x}_n \cdot \boldsymbol{M}$ or as $\lambda \boldsymbol{x}_1 \boldsymbol{x}_2 \cdot \cdot \cdot \boldsymbol{x}_n \boldsymbol{M}$.

We also allow ourselves at any time to introduce abbreviations of the form that a particular symbol α shall stand for a particular sequence of symbols $\boldsymbol{A}$, and indicate the introduction of such an abbreviation by the notation $\alpha \rightarrow \boldsymbol{A}$, to be read, "$\alpha$ stands for $\boldsymbol{A}$."

We introduce at once the following infinite list of abbreviations,

$$1 \rightarrow \lambda ab \cdot a(b),$$
$$2 \rightarrow \lambda ab \cdot a(a(b)),$$
$$3 \rightarrow \lambda ab \cdot a(a(a(b))),$$

and so on, each positive integer in Arabic notation standing for a formula of the form $\lambda ab \cdot a(a(\cdot \cdot \cdot a(b) \cdot \cdot \cdot))$.

The expression $\mathrm{S}^{\boldsymbol{x}}_{\boldsymbol{N}}\boldsymbol{M}\,|$ is used to stand for the result of substituting $\boldsymbol{N}$ for $\boldsymbol{x}$ throughout $\boldsymbol{M}$.

We consider the three following operations on well-formed formulas:

I. *To replace any part* $\lambda \boldsymbol{x}[\boldsymbol{M}]$ *of a formula by* $\lambda \boldsymbol{y}[\mathrm{S}^{\boldsymbol{x}}_{\boldsymbol{y}}\boldsymbol{M}\,|\,]$, *where* $\boldsymbol{y}$ *is a variable which does not occur in* $\boldsymbol{M}$.

II. *To replace any part* $\{\lambda \boldsymbol{x}[\boldsymbol{M}]\}(\boldsymbol{N})$ *of a formula by* $\mathrm{S}^{\boldsymbol{x}}_{\boldsymbol{N}}\boldsymbol{M}\,|$, *provided that the bound variables in* $\boldsymbol{M}$ *are distinct both from* $\boldsymbol{x}$ *and from the free variables in* $\boldsymbol{N}$.

III. *To replace any part* $\mathrm{S}^{\boldsymbol{x}}_{\boldsymbol{N}}\boldsymbol{M}\,|$ (*not immediately following* λ) *of a formula by* $\{\lambda \boldsymbol{x}[\boldsymbol{M}]\}(\boldsymbol{N})$, *provided that the bound variables in* $\boldsymbol{M}$ *are distinct both from* $\boldsymbol{x}$ *and from the free variables in* $\boldsymbol{N}$.

Any finite sequence of these operations is called a *conversion*, and if $\boldsymbol{B}$ is obtainable from $\boldsymbol{A}$ by a conversion we say that $\boldsymbol{A}$ is *convertible* into $\boldsymbol{B}$, or, "$\boldsymbol{A}$ conv $\boldsymbol{B}$." If $\boldsymbol{B}$ is identical with $\boldsymbol{A}$ or is obtainable from $\boldsymbol{A}$ by a single

application of one of the operations I, II, III, we say that $\boldsymbol{A}$ is *immediately convertible* into $\boldsymbol{B}$.

A conversion which contains exactly one application of Operation II, and no application of Operation III, is called a *reduction.*

A formula is said to be *in normal form* if it is well-formed and contains no part of the form $\{\lambda x[M]\}(N)$. And $\boldsymbol{B}$ is said to be a *normal form of* $\boldsymbol{A}$ if $\boldsymbol{B}$ is in normal form and $\boldsymbol{A}$ conv $\boldsymbol{B}$.

The originally given order $a, b, c, \cdots$ of the variables is called their *natural order.* And a formula is said to be *in principal normal form* if it is in normal form, and no variable occurs in it both as a free variable and as a bound variable, and the variables which occur in it immediately following the symbol λ are, when taken in the order in which they occur in the formula, in natural order without repetitions, beginning with a and omitting only such variables as occur in the formula as free variables.[4] The formula $\boldsymbol{B}$ is said to be the *principal normal form of* $\boldsymbol{A}$ if $\boldsymbol{B}$ is in principal normal form and $\boldsymbol{A}$ conv $\boldsymbol{B}$.

Of the three following theorems, proof of the first is immediate, and the second and third have been proved by the present author and J. B. Rosser:[5]

THEOREM I. *If a formula is in normal form, no reduction of it is possible.*

THEOREM II. *If a formula has a normal form, this normal form is unique to within applications of Operation* I, *and any sequence of reductions of the formula must* (*if continued*) *terminate in the normal form.*

THEOREM III. *If a formula has a normal form, every well-formed part of it has a normal form.*

We shall call a function a *function of positive integers* if the range of each independent variable is the class of positive integers and the range of the dependent variable is contained in the class of positive integers. And when it is desired to indicate the number of independent variables we shall speak of a function of one positive integer, a function of two positive integers, and so on. Thus if F is a function of n positive integers, and $a_1, a_2, \cdots, a_n$ are positive integers, then $F(a_1, a_2, \cdots, a_n)$ must be a positive integer.

---

[4] For example, the formulas $\lambda ab \cdot b(a)$ and $\lambda a \cdot a(\lambda c \cdot b(c))$ are in principal normal form, and $\lambda ac \cdot c(a)$, and $\lambda bc \cdot c(b)$, and $\lambda a \cdot a(\lambda a \cdot b(a))$ are in normal form but not in principal normal form. Use of the principal normal form was suggested by S. C. Kleene as a means of avoiding the ambiguity of determination of the normal form of a formula, which is troublesome in certain connections.

Observe that the formulas $1, 2, 3, \cdots$ are all in principal normal form.

[5] Alonzo Church and J. B. Rosser, "Some properties of conversion," forthcoming (abstract in *Bulletin of the American Mathematical Society*, vol. 41, p. 332).

A function F of one positive integer is said to be *λ-definable* if it is possible to find a formula $\boldsymbol{F}$ such that, if $F(m) = r$ and $\boldsymbol{m}$ and $\boldsymbol{r}$ are the formulas for which the positive integers m and r (written in Arabic notation) stand according to our abbreviations introduced above, then $\{\boldsymbol{F}\}(\boldsymbol{m})$ conv $\boldsymbol{r}$.

Similarly, a function F of two positive integers is said to be λ-definable if it is possible to find a formula $\boldsymbol{F}$ such that, whenever $F(m, n) = r$, the formula $\{\boldsymbol{F}\}(\boldsymbol{m}, \boldsymbol{n})$ is convertible into $\boldsymbol{r}$ (m, n, r being positive integers and $\boldsymbol{m}$, $\boldsymbol{n}$, $\boldsymbol{r}$ the corresponding formulas). And so on for functions of three or more positive integers.[6]

It is clear that, in the case of any λ-definable function of positive integers, the process of reduction of formulas to normal form provides an algorithm for the effective calculation of particular values of the function.

3. The Gödel representation of a formula. Adapting to the formal notation just described a device which is due to Gödel,[7] we associate with every formula a positive integer to represent it, as follows. To each of the symbols {, (, [we let correspond the number 11, to each of the symbols },),] the number 13, to the symbol λ the number 1, and to the variables a, b, c, $\cdots$ the prime numbers 17, 19, 23, $\cdots$ respectively. And with a formula which is composed of the n symbols $\tau_1, \tau_2, \cdots, \tau_n$ in order we associate the number $2^{t_1}3^{t_2}\cdots p_n^{t_n}$, where t_i is the number corresponding to the symbol τ_i, and where p_n stands for the n-th prime number.

This number $2^{t_1}3^{t_2}\cdots p_n^{t_n}$ will be called the *Gödel representation* of the formula $\tau_1\tau_2\cdots\tau_n$.

Two distinct formulas may sometimes have the same Gödel representation, because the numbers 11 and 13 each correspond to three different symbols, but it is readily proved that *no two distinct well-formed formulas can have the same Gödel representation.* It is clear, moreover, that there is an effective method by which, given any formula, its Gödel representation can be calculated; and likewise that there is an effective method by which, given any positive integer, it is possible to determine whether it is the Gödel representation of a well-formed formula and, if it is, to obtain that formula.

In this connection the Gödel representation plays a rôle similar to that

---

[6] Cf. S. C. Kleene, "A theory of positive integers in formal logic," *American Journal of Mathematics*, vol. 57 (1935), pp. 153-173 and 219-244, where the λ-definability of a number of familiar functions of positive integers, and of a number of important general classes of functions, is established. Kleene uses the term *definable*, or *formally definable*, in the sense in which we are here using *λ-definable*.

[7] Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I," *Monatshefte für Mathematik und Physik*, vol. 38 (1931), pp. 173-198.

of the matrix of incidence in combinatorial topology (cf. § 1 above). For there is, in the theory of well-formed formulas, an important class of problems, each of which is equivalent to a problem of elementary number theory obtainable by means of the Gödel representation.[8]

4. Recursive functions. We define a class of expressions, which we shall call *elementary expressions*, and which involve, besides parentheses and commas, the symbols 1, S, an infinite set of numerical variables $x, y, z, \cdots$, and, for each positive integer n, an infinite set $f_n, g_n, h_n, \cdots$ of functional variables with subscript n. This definition is by induction as follows. The symbol 1 or any numerical variable, standing alone, is an elementary expression. If A is an elementary expression, then $S(A)$ is an elementary expression. If $A_1, A_2, \cdots, A_n$ are elementary expressions and f_n is any functional variable with subscript n, then $f_n(A_1, A_2, \cdots, A_n)$ is an elementary expression.

The particular elementary expressions 1, $S(1)$, $S(S(1)), \cdots$ are called *numerals*. And the positive integers $1, 2, 3, \cdots$ are said to correspond to the numerals 1, $S(1)$, $S(S(1)), \cdots$.

An expression of the form $A = B$, where A and B are elementary expressions, is called an *elementary equation*.

The *derived equations* of a set E of elementary equations are defined by induction as follows. The equations of E themselves are derived equations. If $A = B$ is a derived equation containing a numerical variable x, then the result of substituting a particular numeral for all the occurrences of x in $A = B$ is a derived equation. If $A = B$ is a derived equation containing an elementary expression C (as part of either A or B), and if either $C = D$ or $D = C$ is a derived equation, then the result of substituting D for a particular occurrence of C in $A = B$ is a derived equation.

Suppose that no derived equation of a certain finite set E of elementary equations has the form $k = l$ where k and l are different numerals, that the functional variables which occur in E are $f_{n_1}{}^1, f_{n_2}{}^2, \cdots, f_{n_r}{}^r$ with subscripts $n_1, n_2, \cdots, n_r$ respectively, and that, for every value of i from 1 to r inclusive, and for every set of numerals $k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i$, there exists a unique numeral k^i such that $f_{n_i}{}^i(k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i) = k^i$ is a derived equation of E. And let $F^1, F^2, \cdots, F^r$ be the functions of positive integers defined by the con-

[8] This is merely a special case of the now familiar remark that, in view of the Gödel representation and the ideas associated with it, symbolic logic in general can be regarded, mathematically, as a branch of elementary number theory. This remark is essentially due to Hilbert (cf. for example, *Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg*, 1904, p. 185; also Paul Bernays in *Die Naturwissenschaften*, vol. 10 (1922), pp. 97 and 98) but is most clearly formulated in terms of the Gödel representation.

dition that, in all cases, $F^i(m_1^i, m_2^i, \cdots, m_{n_i}^i)$ shall be equal to m^i, where $m_1^i, m_2^i, \cdots, m_{n_i}^i$, and m^i are the positive integers which correspond to the numerals $k_1^i, k_2^i, \cdots, k_{n_i}^i$, and k^i respectively. Then the set of equations E is said to *define*, or to be a set of *recursion equations* for, any one of the functions F^i, and the functional variable $f_{n_i}^i$ is said to *denote* the function F^i.

A function of positive integers for which a set of recursion equations can be given is said to be *recursive*.[9]

It is clear that for any recursive function of positive integers there exists an algorithm using which any required particular value of the function can be effectively calculated. For the derived equations of the set of recursion equations E are effectively enumerable, and the algorithm for the calculation of particular values of a function F^i, denoted by a functional variable $f_{n_i}^i$, consists in carrying out the enumeration of the derived equations of E until the required particular equation of the form $f_{n_i}^i(k_1^i, k_2^i, \cdots, k_{n_i}^i) = k^i$ is found.[10]

We call an infinite sequence of positive integers recursive if the function F such that $F(n)$ is the n-th term of the sequence is recursive.

We call a propositional function of positive integers recursive if the function whose value is 2 or 1, according to whether the propositional function is true or false, is recursive. By a recursive property of positive integers we shall mean a recursive propositional function of one positive integer, and by a recursive relation between positive integers we shall mean a recursive propositional function of two or more positive integers.

---

[9] This definition is closely related to, and was suggested by, a definition of recursive functions which was proposed by Kurt Gödel, in lectures at Princeton, N. J., 1934, and credited by him in part to an unpublished suggestion of Jacques Herbrand. The principal features in which the present definition of recursiveness differs from Gödel's are due to S. C. Kleene.

In a forthcoming paper by Kleene to be entitled, "General recursive functions of natural numbers," (abstract in *Bulletin of the American Mathematical Society*, vol. 41), several definitions of recursiveness will be discussed and equivalences among them obtained. In particular, it follows readily from Kleene's results in that paper that every function recursive in the present sense is also recursive in the sense of Gödel (1934) and conversely.

[10] The reader may object that this algorithm cannot be held to provide an effective calculation of the required particular value of F^i unless the proof is constructive that the required equation $f_{n_i}^i(k_1^i, k_2^i, \cdots, k_{n_i}^i) = k^i$ will ultimately be found. But if so this merely means that he should take the existential quantifier which appears in our definition of a set of recursion equations in a constructive sense. What the criterion of constructiveness shall be is left to the reader.

The same remark applies in connection with the existence of an algorithm for calculating the values of a λ-definable function of positive integers.

A function F, for which the range of the dependent variable is contained in the class of positive integers and the range of the independent variable, or of each independent variable, is a subset (not necessarily the whole) of the class of positive integers, will be called *potentially recursive*, if it is possible to find a recursive function F' of positive integers (for which the range of the independent variable, or of each independent variable, is the whole of the class of positive integers), such that the value of F' agrees with the value of F in all cases where the latter is defined.

By an *operation on* well-formed formulas we shall mean a function for which the range of the dependent variable is contained in the class of well-formed formulas and the range of the independent variable, or of each independent variable, is the whole class of well-formed formulas. And we call such an operation recursive if the corresponding function obtained by replacing all formulas by their Gödel representations is potentially recursive.

Similarly any function for which the range of the dependent variable is contained either in the class of positive integers or in the class of well-formed formulas, and for which the range of each independent variable is identical either with the class of positive integers or with the class of well-formed formulas (allowing the case that some of the ranges are identical with one class and some with the other), will be said to be recursive if the corresponding function obtained by replacing all formulas by their Gödel representations is potentially recursive. We call an infinite sequence of well-formed formulas recursive if the corresponding infinite sequence of Gödel representations is recursive. And we call a property of, or relation between, well-formed formulas recursive if the corresponding property of, or relation between, their Gödel representations is potentially recursive. A set of well-formed formulas is said to be recursively enumerable if there exists a recursive infinite sequence which consists entirely of formulas of the set and contains every formula of the set at least once.[11]

In terms of the notion of recursiveness we may also define a *proposition of elementary number theory*, by induction as follows. If ϕ is a recursive propositional function of n positive integers (defined by giving a particular set of recursion equations for the corresponding function whose values are 2 and 1) and if $x_1, x_2, \cdots, x_n$ are variables which take on positive integers as values, then $\phi(x_1, x_2, \cdots, x_n)$ is a proposition of elementary number theory. If P is a proposition of elementary number theory involving x as a free

[11] It can be shown, in view of Theorem V below, that, if an infinite set of formulas is recursively enumerable in this sense, it is also recursively enumerable in the sense that there exists a recursive infinite sequence which consists entirely of formulas of the set and contains every formula of the set exactly once.

variable, then the result of substituting a particular positive integer for all occurrences of x as a free variable in P is a proposition of elementary number theory, and $(x)P$ and $(\exists x)P$ are propositions of elementary number theory, where (x) and $(\exists x)$ are respectively the universal and existential quantifiers of x over the class of positive integers.

It is then readily seen that the negation of a proposition of elementary number theory or the logical product or the logical sum of two propositions of elementary number theory is equivalent, in a simple way, to another proposition of elementary number theory.

5. Recursiveness of the Kleene μ-function. We prove two theorems which establish the recursiveness of certain functions which are definable in words by means of the phrase, "The least positive integer such that," or, "The n-th positive integer such that."

THEOREM IV. *If F is a recursive function of two positive integers, and if for every positive integer x there exists a positive integer y such that $F(x, y) > 1$, then the function F^*, such that, for every positive integer x, $F^*(x)$ is equal to the least positive integer y for which $F(x, y) > 1$, is recursive.*

For a set of recursion equations for F^* consists of the recursion equations for F together with the equations,

$$
\begin{aligned}
&i_2(1, 2) = 2, && g_2(x, 1) = i_2(f_2(x, 1), 2),\\
&i_2(S(x), 2) = 1, && g_2(x, S(y)) = i_2(f_2(x, S(y)), g_2(x, y)),\\
&i_2(x, 1) = 3, && h_2(S(x), y) = x,\\
&i_2(x, S(S(y))) = 3, && h_2(g_2(x, y), x) = j_2(g_2(x, y), y),\\
&j_2(1, y) = y, && f_1(x) = h_2(1, x),\\
&j_2(S(x), y) = x,
\end{aligned}
$$

where the functional variables f_2 and f_1 denote the functions F and F^* respectively, and 2 and 3 are abbreviations for $S(1)$ and $S(S(1))$ respectively.[12]

THEOREM V. *If F is a recursive function of one positive integer, and if there exist an infinite number of positive integers x for which $F(x) > 1$, then the function F^0, such that, for every positive integer n, $F^0(n)$ is equal to the n-th positive integer x (in order of increasing magnitude) for which $F(x) > 1$, is recursive.*

[12] Since this result was obtained, it has been pointed out to the author by S. C. Kleene that it can be proved more simply by using the methods of the latter in *American Journal of Mathematics*, vol. 57 (1935), p. 231 *et seq.* His proof will be given in his forthcoming paper already referred to.

For a set of recursion equations for F^0 consists of the recursion equations for F together with the equations,

$$g_2(1, y) = g_2(f_1(S(y)), S(y)),$$
$$g_2(S(x), y) = y,$$
$$g_1(1) = k,$$
$$g_1(S(y)) = g_2(1, g_1(y)),$$

where the functional variables g_1 and f_1 denote the functions F^0 and F respectively, and where k is the numeral to which corresponds the least positive integer x for which $F(x) > 1$.[13]

6. Recursiveness of certain functions of formulas. We list now a number of theorems which will be proved in detail in a forthcoming paper by S. C. Kleene [14] or follow immediately from considerations there given. We omit proofs here, except for brief indications in some instances.

Our statement of the theorems and our notation differ from Kleene's in that we employ the set of positive integers $(1, 2, 3, \cdots)$ in the rôle in which he employs the set of natural numbers $(0, 1, 2, \cdots)$. This difference is, of course, unessential. We have selected what is, from some points of view, the less natural alternative, in order to preserve the convenience and naturalness of the identification of the formula $\lambda ab \cdot a(b)$ with 1 rather than with 0.

THEOREM VI. *The property of a positive integer, that there exists a well-formed formula of which it is the Gödel representation is recursive.*

THEOREM VII. *The set of well-formed formulas is recursively enumerable.*

This follows from Theorems V and VI.

THEOREM VIII. *The function of two variables, whose value, when taken of the well-formed formulas* $\mathbf{F}$ *and* $\mathbf{X}$, *is the formula* $\{\mathbf{F}\}(\mathbf{X})$, *is recursive.*

THEOREM IX. *The function, whose value for each of the positive integers* $1, 2, 3, \cdots$ *is the corresponding formula* $1, 2, 3, \cdots$, *is recursive.*

THEOREM X. *A function, whose value for each of the formulas* $1, 2, 3, \cdots$ *is the corresponding positive integer, and whose value for other well-formed formulas is a fixed positive integer, is recursive. Likewise the function, whose value for each of the formulas* $1, 2, 3, \cdots$ *is the corresponding positive integer*

---

[13] This proof is due to Kleene.

[14] S. C. Kleene, "λ-definability and recursiveness," forthcoming (abstract in *Bulletin of the American Mathematical Society*, vol. 41). In connection with many of the theorems listed, see also Kurt Gödel, *Monatshefte für Mathematik und Physik*, vol. 38 (1931), p. 181 *et seq.*, observing that every function which is recursive in the sense in which the word is there used by Gödel is also recursive in the present more general sense.

plus one, and whose value for other well-formed formulas is the positive integer 1, *is recursive.*

THEOREM XI. *The relation of immediate convertibility, between well-formed formulas, is recursive.*

THEOREM XII. *It is possible to associate simultaneously with every well-formed formula an enumeration of the formulas obtainable from it by conversion, in such a way that the function of two variables, whose value, when taken of a well-formed formula* **A** *and a positive integer* n, *is the* n*-th formula in the enumeration of the formulas obtainable from* **A** *by conversion, is recursive.*

THEOREM XIII. *The property of a well-formed formula, that it is in principal normal form, is recursive.*

THEOREM XIV. *The set of well-formed formulas which are in principal normal form is recursively enumerable.*

This follows from Theorems V, VII, XIII.

THEOREM XV. *The set of well-formed formulas which have a normal form is recursively enumerable.*[15]

For by Theorems XII and XIV this set can be arranged in an infinite square array which is recursively defined (i. e. defined by a recursive function of two variables). And the familiar process by which this square array is reduced to a single infinite sequence is recursive (i. e. can be expressed by means of recursive functions).

THEOREM XVI. *Every recursive function of positive integers is* λ*-definable.*[16]

THEOREM XVII. *Every* λ*-definable function of positive integers is recursive.*[17]

For functions of one positive integer this follows from Theorems IX, VIII, XII, XIII, IV, X. For functions of more than one positive integer

[15] This theorem was first proposed by the present author, with the outline of proof here indicated. Details of its proof are due to Kleene and will be given by him in his forthcoming paper, "λ-definability and recursiveness."

[16] This theorem can be proved as a straightforward application of the methods introduced by Kleene in the *American Journal of Mathematics* (*loc. cit.*). In the form here given it was first obtained by Kleene. The related result had previously been obtained by J. B. Rosser that, if we modify the definition of *well-formed* by omitting the requirement that M contain x as a free variable in order that $\lambda x[M]$ be well-formed, then every recursive function of positive integers is λ-definable in the resulting modified sense.

[17] This result was obtained independently by the present author and S. C. Kleene at about the same time.

it follows by the same method, using a generalization of Theorem IV to functions of more than two positive integers.

7. The notion of effective calculability. We now define the notion, already discussed, of an *effectively calculable* function of positive integers by identifying it with the notion of a recursive function of positive integers [18] (or of a λ-definable function of positive integers). This definition is thought to be justified by the considerations which follow, so far as positive justification can ever be obtained for the selection of a formal definition to correspond to an intuitive notion.

It has already been pointed out that, for every function of positive integers which is effectively calculable in the sense just defined, there exists an algorithm for the calculation of its values.

Conversely it is true, under the same definition of effective calculability, that every function, an algorithm for the calculation of the values of which exists, is effectively calculable. For example, in the case of a function F of one positive integer, an algorithm consists in a method by which, given any positive integer n, a sequence of expressions (in some notation) $E_{n1}, E_{n2}, \cdots, E_{nr_n}$, can be obtained; where E_{n1} is effectively calculable when n is given; where E_{ni} is effectively calculable when n and the expressions E_{nj}, $j < i$, are given; and where, when n and all the expressions E_{ni} up to and including E_{nr_n} are given, the fact that the algorithm has terminated becomes effectively known and the value of $F(n)$ is effectively calculable. Suppose that we set up a system of Gödel representations for the notation employed in the expressions E_{ni}, and that we then further adopt the method of Gödel of representing a finite sequence of expressions $E_{n1}, E_{n2}, \cdots, E_{ni}$ by the single positive integer $2^{e_{n1}}3^{e_{n2}} \cdots p_i^{e_{ni}}$ where $e_{n1}, e_{n2}, \cdots, e_{ni}$ are respectively the Gödel representations of $E_{n1}, E_{n2}, \cdots, E_{ni}$ (in particular representing a vacuous sequence of expressions by the positive integer 1). Then we may define a function G of two positive integers such that, if x represents the finite sequence $E_{n1}, E_{n2}, \cdots, E_{nk}$, then $G(n, x)$ is equal to the Gödel representation of E_{ni}, where $i = k + 1$, or is equal to 10 if $k = r_n$ (that is if the algorithm has terminated with E_{nk}), and in any other case $G(n, x)$ is equal to 1. And we may define a function H of two positive integers, such that the value of $H(n, x)$ is the same as that of $G(n, x)$, except in the case that $G(n, x) = 10$, in which case $H(n, x) = F(n)$. If the interpretation is allowed that the

[18] The question of the relationship betwen effective calculability and recursiveness (which it is here proposed to answer by identifying the two notions) was raised by Gödel in conversation with the author. The corresponding question of the relationship between effective calculability and λ-definability had previously been proposed by the author independently.

requirement of effective calculability which appears in our description of an algorithm means the effective calculability of the functions G and H,[19] and if we take the effective calculability of G and H to mean recursiveness (λ-definability), then the recursiveness (λ-definability) of F follows by a straightforward argument.

Suppose that we are dealing with some particular system of symbolic logic, which contains a symbol, $=$, for equality of positive integers, a symbol $\{\ \}(\ )$ for the application of a function of one positive integer to its argument, and expressions $1, 2, 3, \cdots$ to stand for the positive integers. The theorems of the system consist of a finite, or enumerably infinite, list of expressions, the *formal axioms*, together with all the expressions obtainable from them by a finite succession of applications of operations chosen out of a given finite, or enumerably infinite, list of operations, the *rules of procedure*. If the system is to serve at all the purposes for which a system of symbolic logic is usually intended, it is necessary that each rule of procedure be an effectively calculable operation, that the complete set of rules of procedure (if infinite) be effectively enumerable, that the complete set of formal axioms (if infinite) be effectively enumerable, and that the relation between a positive integer and the expression which stands for it be effectively determinable. Suppose that we interpret this to mean that, in terms of a system of Gödel representations for the expressions of the logic, each rule of procedure must be a recursive operation,[20] the complete set of rules of procedure must be recursively enumerable (in the sense that there exists a recursive function Φ such that $\Phi(n, x)$ is the representation of the result of applying the n-th rule of procedure to the ordered finite set of formulas represented by x), the complete set of formal axioms must be recursively enumerable, and the relation between a positive integer and the expression which stands for it must be recursive.[21] And let us call a function F of one positive integer[22] *calculable within* the logic if there exists an expression f in the logic such that $\{f\}(\mu) = \nu$ is a theorem when and only when $F(m) = n$ is true, μ and ν being the expressions which stand for the positive integers m and n. Then, since the

[19] If this interpretation or some similar one is not allowed, it is difficult to see how the notion of an algorithm can be given any exact meaning at all.

[20] As a matter of fact, in known systems of symbolic logic, e. g. in that of *Principia Mathematica*, the stronger statement holds, that the relation of *immediate consequence* (*unmittelbare Folge*) is recursive. Cf. Gödel, *loc. cit.*, p. 185. In any case where the relation of immediate consequence is recursive it is possible to find a set of rules of procedure, equivalent to the original ones, such that each rule is a (one-valued) recursive operation, and the complete set of rules is recursively enumerable.

[21] The author is here indebted to Gödel, who, in his 1934 lectures already referred to, proposed substantially these conditions, but in terms of the more restricted notion

complete set of theorems of the logic is recursively enumerable, it follows by Theorem IV above that every function of one positive integer which is calculable within the logic is also effectively calculable (in the sense of our definition).

Thus it is shown that no more general definition of effective calculability than that proposed above can be obtained by either of two methods which naturally suggest themselves (1) by defining a function to be effectively calculable if there exists an algorithm for the calculation of its values (2) by defining a function F (of one positive integer) to be effectively calculable if, for every positive integer m, there exists a positive integer n such that $F(m) = n$ is a provable theorem.

8. Invariants of conversion. The problem naturally suggests itself to find invariants of that transformation of formulas which we have called conversion. The only effectively calculable invariants at present known are the immediately obvious ones (e. g. the set of free variables contained in a formula). Others of importance very probably exist. But we shall prove (in Theorem XIX) that, under the definition of effective calculability proposed in § 7, *no complete set of effectively calculable invariants of conversion exists* (cf. § 1).

The results of Kleene (*American Journal of Mathematics*, 1935) make it clear that, if the problem of finding a complete set of effectively calculable invariants of conversion were solved, most of the familiar unsolved problems of elementary number theory would thereby also be solved. And from Theorem XVI above it follows further that to find a complete set of effectively calculable invariants of conversion would imply the solution of the Entscheidungsproblem for any system of symbolic logic whatever (subject to the very general restrictions of § 7). In the light of this it is hardly surprising that the problem to find such a set of invariants should be unsolvable.

It is to be remembered, however, that, if we consider only the statement of the problem (and ignore things which can be proved about it by more or less lengthy arguments), it appears to be a problem of the same class as the problems of number theory and topology to which it was compared in § 1, having no striking characteristic by which it can be distinguished from them. The temptation is strong to reason by analogy that other important problems of this class may also be unsolvable.

---

of recursiveness which he had employed in 1931, and using the condition that the relation of immediate consequence be recursive instead of the present conditions on the rules of procedure.

[22] We confine ourselves for convenience to the case of functions of one positive integer. The extension to functions of several positive integers is immediate.

LEMMA. *The problem, to find a recursive function of two formulas* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *whose value is* 2 *or* 1 *according as* $\boldsymbol{A}$ *conv* $\boldsymbol{B}$ *or not, is equivalent to the problem, to find a recursive function of one formula* $\boldsymbol{C}$ *whose value is* 2 *or* 1 *according as* $\boldsymbol{C}$ *has a normal form or not.*[23]

For, by Theorem X, the formula $\boldsymbol{a}$ (the formula $\boldsymbol{b}$), which stands for the positive integer which is the Gödel representation of the formula $\boldsymbol{A}$ (the formula $\boldsymbol{B}$), can be expressed as a recursive function of the formula $\boldsymbol{A}$ (the formula $\boldsymbol{B}$). Moreover, by Theorems VI and XII, there exists a recursive function F of two positive integers such that, if m is the Gödel representation of a well-formed formula $\boldsymbol{M}$, then $F(m, n)$ is the Gödel representation of the n-th formula in an enumeration of the formulas obtainable from $\boldsymbol{M}$ by conversion. And, by Theorem XVI, F is λ-definable, by a formula $\mathfrak{f}$. If we define,

$$Z_1 \rightarrow \mathfrak{Q}(\lambda x \cdot x(I), I),$$
$$Z_2 \rightarrow \mathfrak{Q}(\lambda xy \cdot S(x) - y, I),$$

where $\mathfrak{Q}$ is the formula defined by Kleene (*American Journal of Mathematics*, vol. 57 (1935), p. 226), then Z_1 and Z_2 λ-define the functions of one positive integer whose values, for a positive integer n, are the n-th terms respectively of the infinite sequences $1, 1, 2, 1, 2, 3, \cdots$ and $1, 2, 1, 3, 2, 1, \cdots$. By Theorem VIII the formula,

$$\{\lambda xy \cdot \mathfrak{p}(\lambda n \cdot \delta(\mathfrak{f}(x, Z_1(n)), \mathfrak{f}(y, Z_2(n))), 1)\}(\boldsymbol{a}, \boldsymbol{b}),$$

where $\mathfrak{p}$ and δ are defined as by Kleene (*loc. cit.*, p. 173 and p. 231), is a recursive function of $\boldsymbol{A}$ and $\boldsymbol{B}$, and this formula has a normal form if and only if $\boldsymbol{A}$ conv $\boldsymbol{B}$.

Again, by Theorem X, the formula $\boldsymbol{c}$, which stands for the positive integer which is the Gödel representation of the formula $\boldsymbol{C}$, can be expressed as a recursive function of the formula $\boldsymbol{C}$. By Theorems VI and XIII, there exists a recursive function G of one positive integer such that $G(m) = 2$ if m is the Gödel representation of a formula in principal normal form, and $G(m) = 1$ in any other case. And, by Theorem XVI, G is λ-definable, by a formula $\mathfrak{g}$. By Theorem VIII the formula,

$$\{\lambda x \cdot \mathfrak{p}(\lambda n \cdot \mathfrak{g}(\mathfrak{f}(x, n), 1, 1))\}(\boldsymbol{c})$$

[23] These two problems, in the forms, (1) to find an effective method of determining of any two formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ whether $\boldsymbol{A}$ conv $\boldsymbol{B}$, (2) to find an effective method of determining of any formula $\boldsymbol{C}$ whether it has a normal form, were both proposed by Kleene to the author, in the course of a discussion of the properties of the $\mathfrak{p}$-function, about 1932. Some attempts towards solution of (1) by means of numerical invariants were actually made by Kleene at about that time.

where $\mathfrak{f}$ is the formula $\mathfrak{f}$ used in the preceding paragraph, is a recursive function of $\boldsymbol{C}$, and this formula is convertible into the formula 1 if and only if $\boldsymbol{C}$ has a normal form.

Thus we have proved that a formula $\boldsymbol{C}$ can be found as a recursive function of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$, such that $\boldsymbol{C}$ has a normal form if and only if $\boldsymbol{A}$ conv $\boldsymbol{B}$; and that a formula $\boldsymbol{A}$ can be found as a recursive function of a formula $\boldsymbol{C}$, such that $\boldsymbol{A}$ conv 1 if and only if $\boldsymbol{C}$ has a normal form. From this the lemma follows.

THEOREM XVIII. *There is no recursive function of a formula* $\boldsymbol{C}$, *whose value is* 2 *or* 1 *according as* $\boldsymbol{C}$ *has a normal form or not.*

That is, the property of a well-formed formula, that it has a normal form, is not recursive.

For assume the contrary.

Then there exists a recursive function H of one positive integer such that $H(m) = 2$ if m is the Gödel representation of a formula which has a normal form, and $H(m) = 1$ in any other case. And, by Theorem XVI, H is λ-definable by a formula $\mathfrak{h}$.

By Theorem XV, there exists an enumeration of the well-formed formulas which have a normal form, and a recursive function A of one positive integer such that $A(n)$ is the Gödel representation of the n-th formula in this enumeration. And, by Theorem XVI, A is λ-definable, by a formula $\mathfrak{a}$.

By Theorems VI and VIII, there exists a recursive function B of two positive integers such that, if m and n are Gödel representations of well-formed formulas $\boldsymbol{M}$ and $\boldsymbol{N}$, then $B(m, n)$ is the Gödel representation of $\{\boldsymbol{M}\}(\boldsymbol{N})$. And, by Theorem XVI, B is λ-definable, by a formula $\mathfrak{b}$.

By Theorems VI and X, there exists a recursive function C of one positive integer such that, if m is the Gödel representation of one of the formulas $1, 2, 3, \cdots$, then $C(m)$ is the corresponding positive integer plus one, and in any other case $C(m) = 1$. And, by Theorem XVI, C is λ-definable, by a formula $\mathfrak{c}$.

By Theorem IX there exists a recursive function Z^{-1} of one positive integer, whose value for each of the positive integers $1, 2, 3, \cdots$ is the Gödel representation of the corresponding formula $1, 2, 3, \cdots$. And, by Theorem XVI, Z^{-1} is λ-definable, by a formula $\mathfrak{z}$.

Let $\mathfrak{f}$ and $\mathfrak{g}$ be the formulas $\mathfrak{f}$ and $\mathfrak{g}$ used in the proof of the Lemma. By Kleene 15III Cor. (*loc. cit.*, p. 220), a formula $\mathfrak{d}$ can be found such that,

$$\mathfrak{d}(1) \text{ conv } \lambda x \cdot x(1)$$
$$\mathfrak{d}(2) \text{ conv } \lambda u \cdot \mathfrak{c}(\mathfrak{f}(u, \mathfrak{p}(\lambda m \cdot \mathfrak{g}(\mathfrak{f}(u, m)), 1))).$$

We define,

$$\mathfrak{e} \to \lambda n \cdot \mathfrak{d}(\mathfrak{h}(\mathfrak{b}(\mathfrak{a}(n), \mathfrak{z}(n))), \mathfrak{b}(\mathfrak{a}(n), \mathfrak{z}(n))).$$

Then if $\boldsymbol{n}$ is one of the formulas $1, 2, 3, \cdots$, $\mathfrak{e}(\boldsymbol{n})$ is convertible into one of the formulas $1, 2, 3, \cdots$ in accordance with the following rules: (1) if $\mathfrak{b}(\mathfrak{a}(\boldsymbol{n}), \mathfrak{z}(\boldsymbol{n}))$ conv a formula which stands for the Gödel representation of a formula which has no normal form, $\mathfrak{e}(\boldsymbol{n})$ conv 1, (2) if $\mathfrak{b}(\mathfrak{a}(\boldsymbol{n}), \mathfrak{z}(\boldsymbol{n}))$ conv a formula which stands for the Gödel representation of a formula which has a principal normal form which is not one of the formulas $1, 2, 3, \cdots$, $\mathfrak{e}(\boldsymbol{n})$ conv 1, (3) if $\mathfrak{b}(\mathfrak{a}(\boldsymbol{n}), \mathfrak{z}(\boldsymbol{n}))$ conv a formula which stands for the Gödel representation of a formula which has a principal normal form which is one of the formulas $1, 2, 3, \cdots, \mathfrak{e}(\boldsymbol{n})$ conv the next following formula in the list $1, 2, 3, \cdots$.

By Theorem III, since $\mathfrak{e}(1)$ has a normal form, the formula $\mathfrak{e}$ has a normal form. Let $\mathfrak{E}$ be the formula which stands for the Gödel representation of $\mathfrak{e}$. Then, if $\boldsymbol{n}$ is any one of the formulas $1, 2, 3, \cdots$, $\mathfrak{E}$ is not convertible into the formula $\mathfrak{a}(\boldsymbol{n})$, because $\mathfrak{b}(\mathfrak{E}, \mathfrak{z}(\boldsymbol{n}))$ is, by the definition of $\mathfrak{b}$, convertible into the formula which stands for the Gödel representation of $\mathfrak{e}(\boldsymbol{n})$, while $\mathfrak{b}(\mathfrak{a}(\boldsymbol{n}), \mathfrak{z}(\boldsymbol{n}))$ is, by the preceding paragraph, convertible into the formula stands for the Gödel representation of a formula definitely not convertible into $\mathfrak{e}(\boldsymbol{n})$ (Theorem II). But, by our definition of $\mathfrak{a}$, it must be true of one of the formulas $\boldsymbol{n}$ in the list $1, 2, 3, \cdots$ that $\mathfrak{a}(\boldsymbol{n})$ conv $\mathfrak{E}$.

Thus, since our assumption to the contrary has led to a contradiction, the theorem must be true.

In order to present the essential ideas without any attempt at exact statement, the preceding proof may be outlined as follows. We are to deduce a contradiction from the assumption that it is effectively determinable of every well-formed formula whether or not it has a normal form. If this assumption holds, it is effectively determinable of every well-formed formula whether or not it is convertible into one of the formulas $1, 2, 3, \cdots$; for, given a well-formed formula $\boldsymbol{R}$, we can first determine whether or not it has a normal form, and if it has we can obtain the principal normal form by enumerating the formulas into which $\boldsymbol{R}$ is convertible (Theorem XII) and picking out the first formula in principal normal form which occurs in the enumeration, and we can then determine whether the principal normal form is one of the formulas $1, 2, 3, \cdots$. Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \cdots$ be an effective enumeration of the well-formed formulas which have a normal form (Theorem XV). Let E be a function of one positive integer, defined by the rule that, where $\boldsymbol{m}$ and $\boldsymbol{n}$ are the formulas which stand for the positive integers m and n respectively, $E(n) = 1$ if $\{\boldsymbol{A}_n\}(\boldsymbol{n})$ is not convertible into one of the formulas $1, 2, 3, \cdots$, and $E(n) = m + 1$ if $\{\boldsymbol{A}_n\}(\boldsymbol{n})$ conv $\boldsymbol{m}$ and $\boldsymbol{m}$ is one of the formulas $1, 2, 3, \cdots$. The function E is effectively calculable and is there-

fore λ-definable, by a formula $\mathfrak{e}$. The formula $\mathfrak{e}$ has a normal form, since $\mathfrak{e}(1)$ has a normal form. But $\mathfrak{e}$ is not any one of the formulas $A_1, A_2, A_3, \cdots$, because, for every n, $\mathfrak{e}(n)$ is a formula which is not convertible into $\{A_n\}(n)$. And this contradicts the property of the enumeration $A_1, A_2, A_3, \cdots$ that it contains all well-formed formulas which have a normal form.

COROLLARY 1. *The set of well-formed formulas which have no normal form is not recursively enumerable.*[24]

For, to outline the argument, the set of well-formed formulas which have a normal form is recursively enumerable, by Theorem XV. If the set of those which do not have a normal form were aslo recursively enumerable, it would be possible to tell effectively of any well-formed formula whether it had a normal form, by the process of searching through the two enumerations until it was found in one or the other. This, however, is contrary to Theorem XVIII.

This corollary gives us an example of an effectively enumerable set (the set of well-formed formulas) which is divided into two non-overlapping subsets of which one is effectively enumerable and the other not. Indeed, in view of the difficulty of attaching any reasonable meaning to the assertion that a set is enumerable but not effectively enumerable, it may even be permissible to go a step further and say that here is an example of an enumerable set which is divided into two non-overlapping subsets of which one is enumerable and the other non-enumerable.[25]

COROLLARY 2. *Let a function F of one positive integer be defined by the rule that $F(n)$ shall equal 2 or 1 according as n is or is not the Gödel representation of a formula which has a normal form. Then F (if its definition be admitted as valid at all) is an example of a non-recursive function of positive integers.*[26]

This follows at once from Theorem XVIII.

---

[24] This corollary was proposed by J. B. Rosser.

The outline of proof here given for it is open to the objection, recently called to the author's attention by Paul Bernays, that it ostensibly requires a non-constructive use of the principle of excluded middle. This objection is met by a revision of the proof, the revised proof to consist in taking any recursive enumeration of formulas which have no normal form and showing that this enumeration is not a complete enumeration of such formulas, by constructing a formula $\mathfrak{e}(n)$ such that (1) the supposition that $\mathfrak{e}(n)$ occurs in the enumeration leads to contradiction (2) the supposition that $\mathfrak{e}(n)$ has a normal form leads to contradiction.

[25] Cf. the remarks of the author in *The American Mathematical Monthly*, vol. 41 (1934), pp. 356-361.

[26] Other examples of non-recursive functions have since been obtained by S. C. Kleene in a different connection. See his forthcoming paper, "General recursive functions of natural numbers."

Consider the infinite sequence of positive integers, $F(1), F(2), F(3), \cdots$. It is impossible to specify effectively a method by which, given any n, the n-th term of this sequence could be calculated. But it is also impossible ever to select a particular term of this sequence and prove about that term that its value cannot be calculated (because of the obvious theorem that if this sequence has terms whose values cannot be calculated then the value of each of those terms 1). Therefore it is natural to raise the question whether, in spite of the fact that there is no systematic method of effectively calculating the terms of this sequence, it might not be true of each term individually that there existed a method of calculating its value. To this question perhaps the best answer is that the question itself has no meaning, on the ground that the universal quantifier which it contains is intended to express a mere infinite succession of accidents rather than anything systematic.

There is in consequence some room for doubt whether the assertion that the function F exists can be given a reasonable meaning.

THEOREM XIX. *There is no recursive function of two formulas* $\boldsymbol{A}$ *and* $\boldsymbol{B}$, *whose value is* 2 *or* 1 *according as* $\boldsymbol{A}$ *conv* $\boldsymbol{B}$ *or not.*

This follows at once from Theorem XVIII and the Lemma preceding it.

As a corollary of Theorem XIX, it follows that the Entscheidungsproblem is unsolvable in the case of any system of symbolic logic which is ω-consistent (ω-widerspruchsfrei) in the sense of Gödel (*loc. cit.*, p. 187) and is strong enough to allow certain comparatively simple methods of definition and proof. For in any such system the proposition will be expressible about two positive integers a and b that they are Gödel representations of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A}$ is immediately convertible into $\boldsymbol{B}$. Hence, utilizing the fact that a conversion is a finite sequence of immediate conversions, the proposition $\Psi(a, b)$ will be expressible that a and b are Gödel representations of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A}$ conv $\boldsymbol{B}$. Moreover if $\boldsymbol{A}$ conv $\boldsymbol{B}$, and a and b are the Gödel representations of $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively, the proposition $\Psi(a, b)$ will be provable in the system, by a proof which amounts to exhibiting, in terms of Gödel representations, a particular finite sequence of immediate conversions, leading from $\boldsymbol{A}$ to $\boldsymbol{B}$; and if $\boldsymbol{A}$ is not convertible into $\boldsymbol{B}$, the ω-consistency of the system means that $\Psi(a, b)$ will not be provable. If the Entscheidungsproblem for the system were solved, there would be a means of determining effectively of every proposition $\Psi(a, b)$ whether it was provable, and hence a means of determining effectively of every pair of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ whether $\boldsymbol{A}$ conv $\boldsymbol{B}$, contrary to Theorem XIX.

In particular, if the system of *Principia Mathematica* be ω-consistent, its Entscheidungsproblem is unsolvable.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

ASYMPTOTIC SOLUTIONS OF CERTAIN ORDINARY DIFFERENTIAL EQUATIONS ASSOCIATED WITH MULTIPLE ROOTS OF THE CHARACTERISTIC EQUATION.

By HUGH L. TURRITTIN.

Introduction. A linear differential equation

$$(1)\quad \mathfrak{L}(y) \equiv \frac{d^n y}{dx^n} + \rho^r P_1(x,\rho)\frac{d^{n-1}y}{dx^{n-1}} + \rho^{2r}P_2(x,\rho)\frac{d^{n-2}y}{dx^{n-2}} + \cdots + \rho^{nr}P_n(x,\rho)y = 0$$

with continuous coefficients that may be represented by convergent series

$$(2)\quad P_i(x,\rho) = \sum_{j=0}^{\infty} P_{i,j}(x)\rho^{-j}, \quad (i = 1, 2, \cdots, n;\ a \leqq x \leqq b;\ |\rho| > R_1)$$

has under certain well known conditions [1] a fundamental set of solutions of the form

$$(3)\quad y_i(x,\rho) = \left(\sum_{j=0}^{\zeta} y_{ij}(x)\rho^{-j} + \frac{\mathcal{E}(x,\rho)}{\rho^{\zeta+1}}\right)\exp\left[\int_a^x \Omega_i(t,\rho)\,dt\right]$$
$$(a \leqq x \leqq b;\ |\rho| > R_3 > R_1;\ \gamma \leqq \arg.\,\rho \leqq \theta)$$

where $\mathcal{E}$ is a generic symbol for a uniformly bounded function. The particular form of the solutions as exhibited in (3) has been derived in the classical theory subject to the hypothesis that the characteristic equation possesses distinct roots throughout the fundamental interval (a, b). The object of the present paper is to generalize the theory and dispense in part with the above mentioned hypothesis. Schlesinger stated [2] and the analysis below will prove that, if multiple characteristic roots occur, it is necessary to modify the right-hand member of (3) so as to include fractional, as well as full, powers of the parameter ρ.

[1] The development of the theory concerned with solutions of type (3) may be traced through the works of H. Poincaré, *Acta Mathematica*, vol. 8 (1886), pp. 295-344; J. Horn, *Mathematische Annalen*, vol. 52 (1899), pp. 271-292; L. Schlesinger, *Mathematische Annalen*, vol. 63 (1907), pp. 277-300; G. D. Birkhoff, *Transactions of the American Mathematical Society*, vol. 9 (1908), pp. 219-231; O. Perron, *Sitzungsberichte der Heidelberger Akamedieder Wissenschaften, Mathematik Naturwissenschaften*, Klasse 1919 (A. 6); J. Tamarkin, *Thesis*, Petrograd (1917), and *Mathematische Zeitschrift*, vol. 27 (1928), pp. 1-54.

[2] L. Schlesinger, *op. cit.*, p. 282.

The first part of the demonstration is devoted to the construction of n formal series solutions. The existence proof proper, which then follows, begins with the study of a differential equation satisfied by n functions $z_i(x, \rho)$ created from the formal series by cutting them short at some particular term. The asymptotic relation of these functions to actual solutions of the differential equation (1) is established by means of a convenient integral equation and certain related inequalities.

The characteristic equation. Previous investigators have shown that to each simple root of the characteristic equation

$$(4) \qquad F(\phi) \equiv \phi^n + P_{1,0}(x)\phi^{n-1} + P_{2,0}(x)\phi^{n-2} + \cdots + P_{n,0}(x) = 0$$

there may be made to correspond a formal solution of differential equation (1). In order to procure n independent expansions when multiple roots are admitted, it would therefore seem necessary that with each k-fold characteristic root $\phi(x)$ there must be associated k formal series. This will essentially be done, not, however, with respect to the single characteristic equation (4), but with reference to a sequence of auxiliary characteristic equations.

Throughout the text it will be convenient to indicate derivatives by superscripts in parentheses. In particular reference will be made to the function $F^{(j)}(\phi)$, the j-th derivative with respect to ϕ of the left member of the characteristic equation (4).

Hypotheses. It will be understood hereafter that a root which is distinct from all others throughout the fundamental interval (a, b) is to be called *simple*; while the term *multiple* will be reserved for all those roots, and only those, which coincide over the entire interval (a, b). It is to be assumed that all roots, whether of equation (4), or of the subsequent auxiliary characteristic equations, are to be, in the sense just defined, either simple or multiple. This paper will not be concerned with roots of a more complicated nature.[3]

Furthermore it is presumed that the functions $P_{i,j}(x)$ in series (2) possess derivatives of all orders.

Later on it will be assumed that certain Wronskians do not vanish. Such assumptions will be indicated by the letter $(\mathcal{A})$.

If a change of variable is made by setting $y = ve^{-\rho^r \epsilon x}$, where ϵ is a small positive constant, the general form of the differential equation (1) remains unaltered. However, the characteristic equation is affected and the char-

[3] See, however, in this regard R. E. Langer, *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 90-106.

acteristic roots are increased by an amount equal to ϵ. Such a change in the equation can always be employed to avoid zero characteristic roots. It therefore can and will be assumed without imposing an additional restriction that $P_{n,0}(x) \not\equiv 0$.

The first change of variable and the Puiseux polygon. As an initial step toward procuring formal series solutions corresponding to a k-fold characteristic root $\phi(x)$, substitute in (1)

$$y = v \exp[\int_a^x \rho^r \phi(t)\, dt].$$

The differential equation (1) then acquires the new form

$$(5) \qquad v^{(n)} + p_1(x,\rho)v^{(n-1)} + p_2(x,\rho)v^{(n-2)} + \cdots + p_n(x,\rho)v = 0$$

where the leading term in the series expansion of the coefficient $p_i(x,\rho)$ is formally $[\rho^{ir}/(n-i)!]F^{(n-i)}(\phi)$. Due to the k-fold multiplicity of $\phi(x)$, $F^{(n-i)}(\phi) \equiv 0$ if $i = n, n-1, \cdots$ or $n-k+1$. Consequently the coefficients have the structure

$$(6) \qquad p_i(x,\rho) = \rho^{j_i} p_{i,0}(x) + \rho^{j_i} \sum_{h=1}^{\infty} p_{i,h}(x)\rho^{-h} \qquad (i = 1, \cdots, n)$$

with

$$\begin{aligned} &j_i \leqq ir && \text{for} \quad i = 1, 2, \cdots, n-k-1; \\ &j_i = (n-k)r && \text{for} \quad i = n-k, \text{ and} \\ &j_i \leqq ir - 1 && \text{for} \quad i = n-k+1, \cdots, n. \end{aligned}$$

Hereafter the leading term in an expansion of a function in descending powers of ρ will be called the *principal term.*

The principal terms of the series (6) can be advantageously represented by a diagram in the following manner:

Let $p_{i,0}(x)$ correspond to the point with coördinates (i, j_i); join the origin to the point $(n, 0)$ not only by a straight line-segment, but also by a broken line to form a convex polygon with vertices falling on points (i, j_i) and enclosing the greatest possible positive area. The result is an adaptation of the familiar Puiseux polygon.[4] By virtue of the construction the slope s_i of the i-th line-segment (i. e. the slope of the i-th side of the polygon counting from left to right) is less than the slope s_{i-1} of the preceding segment, and the points (i, j_i) will all fall on, or below, the broken line. The significance of the Puiseux polygon lies in the fact that there can be associated with each

[4] V. Puiseux, *Journal de Mathématiques Pures et Appliquées*, vol. 11 (1889), p. 319.

of the line-segments formal solutions of equation (1) in number equal to the length of the projection of the respective segment, or side, on the x-axis.

Case I. Formal solutions pertaining to segments with negative or zero slope are derived as follows:

Let the coördinates of the right terminus of the chain of segments with positive slopes be designated by (m, j_m). Then replace the dependent variable v in equation (5) by the series

$$v_0(x) + v_1(x)/\rho + v_2(x)/\rho^2 + \cdots.$$

Equate the coefficients of the various powers of ρ to zero; and thus derive a sequence of linear differential equations, each of order $n - m$ from which the $v_i(x)$'s may be successively determined. The first differential equation of this sequence is homogeneous, and therefore yields $n - m$ independent evaluations of $v_0(x)$. The other $v_i(x)$'s may be determined by quadratures once the $v_0(x)$'s have been found. Hence there exist for the differential equation (1) $n - m$ formal solutions of the structure

$$(7)\quad (v_{j,0}(x) + v_{j,1}(x)/\rho + \cdots) \exp\Big[\int_a^x \rho^r \phi(x)\,dx\Big] \quad (j = 1, 2, \cdots, n - m).$$

It will be assumed, (A), that the Wronskian

$$(8)\qquad W(v_{1,0}(x), \cdots, v_{n-m,0}(x)) \neq 0;$$

and the hypothesis (8) would be satisfied, it is to be noted, if $p_{m,0}(x)$ were non-vanishing on (a, b).

Case II. There remains for consideration only those segments of positive slope which lie between the points $(n - k, nr - kr)$ and (m, j_m), for the first segment pertains to characteristic roots other than $\phi(x)$. Let attention be fixed upon any particular one of these segments, the σ-th for instance, and let μ_σ denote the length of the projection of the segment upon the x-axis. To obtain formal solutions to the number μ_σ, one, or more, transformations will be necessary. Let

$$(9)\qquad v = u \exp\Big[\int_a^x \lambda^\alpha \psi_1(x)\,dx\Big]$$

with $\lambda^\alpha = \rho^{\alpha/\beta}$, where α/β is a fraction in lowest terms equal to the slope s_σ of the segment in question.

Transformation (9) modifies simultaneously the Puiseux polygon and the differential equation (5). Those segments up to and including the

$(\sigma-1)$-th remain unaltered. The chain of segments to the right of the $(\sigma-1)$-th is, however, replaced by a single segment of slope s_σ. The differential equation (5) is reduced by (9) to a form which, in so far as the last $(\mu_\sigma+1)$ terms are concerned, can be indicated as follows:

$$\cdots+q_0(x,\lambda)u^{(\mu_\sigma)}+q_1(x,\lambda)u^{(\mu_\sigma-1)}+\cdots+q_{\mu_\sigma}(x,\lambda)u=0$$

in which

$$(10)\quad q_j(x,\lambda)=\frac{\lambda^{\alpha_j}G^{(\mu_\sigma-j)}(\psi_1)}{(\mu_\sigma-j)!}+\lambda^{\alpha_j}\sum_{h=1}^{\infty}q_{j,h}(x)\lambda^{-h}\qquad(j=0,1,\cdots,\mu_\sigma)$$

with $G^{(\mu_\sigma-j)}(\psi)$ representing the $(\mu_\sigma-j)$-th derivative of the function

$$(11)\qquad G(\psi)\equiv\psi^{n-\mu_1-\mu_2-\cdots-\mu_\sigma}\{C_0(x)\,\psi^{\mu_\sigma}+\cdots+C_{\mu_\sigma}(x)\}=0$$

which, as is indicated, has been set equal to zero in (11) to form the auxiliary characteristic equation. The coefficients $C_0(x)$ and $C_{\mu_\sigma}(x)$ are equal to those functions $p_{i,0}(x)$ of series (6) which correspond to the vertices at the left and right ends respectively of the σ-th segment in the polygon. The intermediate coefficients of (11) likewise are equal to $p_{i,0}(x)$'s which correspond to other points (i,j_i) that fall on the σ-th segment. These facts make it clear that transformation (9) has in effect shifted the σ-th segment of the original polygon over to the extreme right of the figure.

The portion of the auxiliary characteristic equation (11) which is contained within the braces is a polygon in ψ^β. Hence there are at least β different non-zero roots of the equation (11). If one root has been found, $\beta-1$ others can be computed by multiplying the known one by the various β-th roots of unity. When in the subsequent discussion the function $\psi_1(x)$ which appears in substitution (9) is identified with a particular root of (11), it should be born in mind that β such roots can be considered simultaneously by allowing the β-th roots of unity to be absorbed into the parameter λ.

The function $\psi_1(x)$ of substitution (9), which as yet has not been specified, will be selected as a non-zero κ-fold multiple root of the auxiliary characteristic equation. Such a choice has an immediate effect upon the differential equation causing the leading terms in those series of (10) to vanish for which $j=\mu_\sigma,\mu_\sigma-1,\cdots$ and $\mu_\sigma-\kappa+1$. The effect upon the polygon is even more important to note. One of two things must take place. Either the right-hand portion of the new diagram which corresponds to the old σ-th segment is broken up into at least two smaller line-segments, or the right-hand segment is subject to a decrease in slope. The break-up certainly takes place if $\beta>1$ and the decrease in slope occurs if $\beta=1$.

Since κ formal series are to be associated with ψ_1, we need no longer be concerned with the original σ-th segment, but must concentrate upon those segments between $x = n - \kappa$ and $x = n$, all of which have slopes less than s_σ. On the one hand certain of these segments may have zero or negative slope and to such segments formal series will at once be associated by the method described in Case I, subject of course to a hypothesis $(\mathcal{A})$ that a certain Wronskian, the counterpart of (8), is non-vanishing. On the other hand, if formal solutions are to be associated with those segments of positive slope between $x = n - \kappa$ and $x = n$, further exponential transformations must be made.

The Puiseux polygon which has undergone two changes, one brought about by substitution (9), the other by virtue of a special choice of ψ_1, supplies the data for selecting the next exponential transformation; in fact the power of the parameter which must appear under the integral sign equals the slope of the line segment to which formal solutions are eventually to be associated. For instance, if a segment between $x = n - \kappa$ and $x = n$ has a slope Δ, $(0 < \Delta < s_\sigma)$, the proper substitution is

$$u = w \exp[\int_a^x \rho^\Delta \chi(x)\, dx]. \tag{12}$$

The effect of this transformation can be readily visualized in as much as the effect of (9), a typical transformation, has already been described in detail. Transformation (12) introduces, as a rule, a new and smaller fractional power of ρ into the differential equation and the polygon is modified to the extent that the segment with slope Δ is shifted to the right side of the diagram. Then a second auxiliary characteristic equation is formed by equating to zero the principal term in the series expansion of the coefficient of w in the new linear differential equation. The function $\chi(x)$ is definitely chosen as one of the non-zero secondary auxiliary characteristic roots with the result that the Puiseux polygon is either broken into smaller segments or the slope of the right-hand segment is decreased. Formal solutions are immediately associated with all those segments which have, as a result of the reduction, zero or negative slopes. (This reapplication of Case I calls, in general, for another assumption of type $\mathcal{A}$.) To associate formal solutions with the segments of positive slope still further transformation are required, and hence, *mutatis mutandis*, the whole process is repeated.

A succession of reductions, each involving a decrease in slope, unaccompanied by a break-up of the segments, must give rise eventually to a segment of zero slope and the process is thereby terminated (Case I, $\mathcal{A}$). That the

zero slope is ultimately attained is made evident by reexamining the typical reduction arising from (9) when the appropriate values $\beta = 1$, $\lambda = \rho$, and the integer $\alpha \leq r - 1$ are substituted in (9). The reduction will result in a slope $\leq r - 2$ for the right-hand segment, because no new fractional power of the parameter has been introduced. A second reapplication of the process with a new exponent $\alpha_1 \leq r - 2$ will result in a slope $\leq r - 3$, etc., until with not more than $r - 1$ successive reductions, with β continually equal to one, a zero slope is attained for the right-hand segment of the polygon. On the other hand a series of reductions, accompanied by a succession of break-ups, must finally result in a short segment with unit projection on the x-axis. The auxiliary characteristic equation, which comes from a reapplication of the process to such a segment, is of the first degree and can not possess a multiple root. Therefore from this stage on only decreases in slope can occur when the proper exponential transformations are made. Finally a zero slope is attained and formal solutions procured. This reasoning establishes the fact that there exist n formal solutions for differential equation (1) of the structure

$$(13)\quad \left(z_{i,0}(x) + \frac{z_{i,1}(x)}{\rho^{1/\beta_i}} + \frac{z_{i,2}(x)}{\rho^{2/\beta_i}} + \cdots\right) \exp\left[\int_a^x \Omega_i(t,\rho)\,dt\right] \quad (i = 1, \cdots, n)$$

where $\Omega_i(t,\rho)$ is an abbreviation for the sum

$$\rho^r \Omega_{i,1}(t) + \rho^{r-1/\beta_i}\Omega_{i,2}(t) + \cdots + \rho^{1/\beta_i}\Omega_{i,r\beta_i-1}(t).$$

It is, moreover, allowable to attribute to ρ^{1/β_i} any one of its β_i values. Furthermore, if for some particular value of i, $\Omega_i(t,\rho)$ is different from all other $\Omega_j(t,\rho)$, $i \neq j$, then it follows as a specialization of (8) that

$$(14)\qquad z_{i,0}(x) \not\equiv 0.$$

The related differential equation. In order to obtain an asymptotic relationship between the formal series (13) and actual solutions of the given differential equation (1), a study will be made of the differential equation satisfied by the functions $z_i(x,\rho)$ defined by arbitrarily terminating series (13); i. e.,

$$z_i(x,\rho) = \eta_i(x,\rho) \exp\left[\int_a^x \Omega_i(t,\rho)\,dt\right] \qquad (i = 1, \cdots, n).$$

with

$$\eta_i(x,\rho) = z_{i,0}(x) + z_{i,1}(x)/\rho^{1/\beta_i} + \cdots + z_{i,\zeta\beta_i}(x)/\rho^{\zeta}.$$

These functions $z_i(x,\rho)$ are solutions of the differential equation

$$(15)\qquad \begin{vmatrix} z^{(n)} & \cdot & \cdot & \cdot & z^{(1)} & z \\ z_1^{(n)} & \cdot & \cdot & \cdot & z_1^{(1)} & z_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ z_n^{(n)} & \cdot & \cdot & \cdot & z_n^{(1)} & z_n \end{vmatrix} = 0$$

which has been written as a determinant. This determinant becomes, when expanded in terms of the elements of the first row,

$$(16)\qquad \mathfrak{M}(z) \equiv z^{(n)} + \rho^r Q_1(x,\rho) z^{(n-1)} + \cdots + \rho^{rn} Q_n(x,\rho) z = 0$$

where the coefficients are quotients of two n-th order determinants; i. e.,

$$(17)\qquad Q_i(x,\rho) = (-1)^i A_i(x,\rho)/\rho^{ir} A_0(x,\rho) \qquad (i = 1, \cdots, n)$$

In particular $A_0(x,\rho)$ is the Wronskian $W(z, \cdots, z_n)$.

Since the various functions $z_i(x,\rho)$, which are associated with the introduction of a particular fractional power of ρ, may be permuted among themselves in (15) without essentially changing the differential equation (16), it is evident that the expansions of the determinants $A_i(x,\rho)$ proceed according to decreasing *full* powers. Furthermore these expansions are finite, not infinite series.

Structure of the coefficients $Q_i(x,\rho)$. To make certain that $Q_i(x,\rho)$ can be expanded in convergent power series in ρ^{-1}, it will be sufficient to show that $A_0(x,\rho) \exp[-\int_a^x (\sum_{i=1}^n \Omega_i) dt]$, which is later indicated by $\Lambda A_0(x,\rho)$, is bounded away from zero as $\rho \to \infty$. The element in the i-th row and j-th column of $A_0(x,\rho)$ is

$$d^{n-j} z_i / dx^{n-j} = (d_i/dx + \Omega_i(x,\rho))^{n-j} \eta_i(x,\rho) \exp[\int_a^x \Omega_i dt]$$

$$= \mathfrak{D}_i^{n-j} \eta_i(x,\rho) \exp[\int_a^x \Omega_i dt].$$

The symbolic operators $\mathfrak{D}_i = (d_i/dx + \Omega_i(x,\rho))$ and d_i/dx will be especially useful in ascertaining the character of $A_0(x,\rho)$ as $\rho \to \infty$. It should be understood, however, that, when the differential operator d_i/dx acts upon a function with subscript other than i, it is to be considered equivalent to differentiating a constant. For example this notation permits one first to think of the Wronskian as a Vandermonde determinant; i. e.,

$$\begin{vmatrix} \eta_1 & \cdots & \eta_n \\ \eta_1^{(1)} & \cdots & \eta_n^{(1)} \\ \cdot & \cdots & \cdot \\ \eta_1^{(n-1)} & \cdots & \eta_n^{(n-1)} \end{vmatrix} = \{\prod_{i>j}(d_i/dx - d_j/dx)\}\eta_1\eta_2\cdots\eta_n; \quad \begin{matrix}(i=2,3,\cdots,n;\\ j=1,\cdots,n-1),\end{matrix}$$

and secondly to write

$$(18)\quad \Lambda A_0(x,\rho) = \begin{vmatrix} \mathcal{D}_1^{n-1} & \cdots & \mathcal{D}_1 & 1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ \mathcal{D}_n^{n-1} & \cdots & \mathcal{D}_n & 1 \end{vmatrix} \eta_1\eta_2\cdots\eta_n = \{\prod_{i<j}(\mathcal{D}_i - \mathcal{D}_j)\}\eta_1\eta_2\cdots\eta_n.$$
$$(i=1,\cdots,n-1;\ j=2,\cdots,n),$$

a convenient expression for $\Lambda A_0(x,\rho)$. A glance at the right-hand member of (18) shows, if we consider the principal terms in the expansions of the various factors, that as $\rho\to\infty$ the behavior of $\Lambda A_0(x,\rho)$ is essentially the same as that of the product

$$(19)\qquad \{\prod_{i<j}(\Omega_i - \Omega_j)\}\{\prod_{i<j}(d_i/dx - d_j/dx)\}z_{1,0}z_{2,0}\cdots z_{n,0}$$

where all those combinations of i and j $(i<j)$, for which $\Omega_i - \Omega_j \not\equiv 0$, are to be used in the first brace, and in the second all other combinations of $i<j\leq n$ are to be used. The product (19) and likewise $\Lambda A_0(x,\rho)$ are bounded away from zero for $|\rho|$ sufficiently large due to the fact that in (19) only non-vanishing factors occur, some of these factors are of type $(\Omega_i - \Omega_j)$, some of type (14), and others are non-zero Wronskians of type (8). Not only does the above argument establish the fact that $Q_i(x,\rho)$ can be expanded in convergent series in descending powers of ρ, but it also shows that the functions $z_1(x,\rho),\cdots,z_n(x,\rho)$ are independent.

To demonstrate that $Q_i(x,\rho)$ is uniformly bounded in the neighborhood of $\rho=\infty$, it is enough to show that

$$(20)\qquad Q_i(x,\rho) = \sum_{h=0}^{\infty} Q_{i,h}(x)\rho^{-h} \qquad (i=1,\cdots,n).$$

Let $E_j(\mathcal{D}_1,\cdots,\mathcal{D}_n)$ denote the j-th elementary symmetric function $\Sigma\mathcal{D}_1\mathcal{D}_2\cdots\mathcal{D}_j$ and $E_0=1$; and then, from the fact[5] that the determinant

[5] E. R. Heineman, *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 464-476.

$$\begin{vmatrix} \mathfrak{D}_1^n & \mathfrak{D}_1^{n-1} & \cdots & \mathfrak{D}_1^{j+1} & \mathfrak{D}_1^{j-1} & \cdots & \mathfrak{D}_1 & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \mathfrak{D}_n^n & \mathfrak{D}_n^{n-1} & \cdots & \mathfrak{D}_n^{j+1} & \mathfrak{D}_n^{j-1} & \cdots & \mathfrak{D}_n & 1 \end{vmatrix} = E_{n-j}(\mathfrak{D}_1, \cdots, \mathfrak{D}_n) \prod_{i<h} (\mathfrak{D}_i - \mathfrak{D}_h),$$

$$(j = 0, 1, \cdots, n;\ i = 1, \cdots, n-1;\ h = 2, \cdots, n),$$

it follows that

$$\Lambda A_i(x, \rho) = E_i(\mathfrak{D}_1, \cdots, \mathfrak{D}_n) \prod_{j<h} (\mathfrak{D}_j - \mathfrak{D}_h)\eta_1\eta_2 \cdots \eta_n. \tag{21}$$

$$(i = 1, \cdots, n;\ j = 1, \cdots, n-1;\ h = 2, \cdots, n).$$

If it is taken into account that the power of ρ in the principal term of $E_i(\mathfrak{D}_1, \cdots, \mathfrak{D}_n)$ can not exceed ρ^{ir} and then if formulae (18) and (21) are compared, it will be evident that the quotients (17) must have the structure (20).

A comparison of the related and original differential equation. The discussion of the last section brings out the structural similarity of the given differential equation (1) and the related equation (16). The resemblance will be made more striking by showing (see inequality (24)) that the respective coefficients of these differential equations are to a certain extent identical; i. e.,

$$P_{i,j}(x) \equiv Q_{i,j}(x) \quad \text{for} \quad \begin{cases} i = 1, \cdots, n \\ j = 0, 1, \cdots, \zeta + r. \end{cases}$$

With this end in view consider the n differences

$$\mathfrak{M}(z_i) - \mathfrak{L}(z_i). \tag{22}$$

Since $z_i(x, \rho)$ is a solution of (16)

$$\mathfrak{M}(z_i(x, \rho)) \equiv 0.$$

The substitution of $z_i(x, \rho)$ in (1) must cause a number of terms to vanish by virtue of the fact that $z_i(x, \rho)$ is a portion of a formal series solution; those terms which do not in general vanish may be indicated as follows:

$$\mathfrak{L}(z_i(x, \rho)) \equiv \rho^{\tau_i - \zeta - 1/\beta_i} (f_{i,0}(x) + f_{i,1}(x)/\rho^{1/\beta_i} + \cdots) \exp[\int_a^x \Omega_i dt] \tag{23}$$

$$(i = 1, \cdots, n).$$

The f's are appropriate functions of x, and τ_i designates the power of ρ occurring in the principal term of the expansion of $\prod_j (\Omega_j - \Omega_i)$, where $j \leq n$

ranges over all values such that $\Omega_j \not\equiv \Omega_i$. The differences (22) may be rewritten by utilizing (23) as

$$\sum_{j=1}^{n} \rho^{jr}(P_j - Q_j) z_i^{(n-j)} \equiv \mathfrak{L}(z_i) \qquad (i = 1, \cdots, n).$$

These identities are treated as n simultaneous equations defining the unknowns $(P_1 - Q_1), \cdots, (P_n - Q_n)$. Each of these unknowns can be expressed by means of Cramer's rule as a quotient of two determinants. The symbolic operators $\mathfrak{D}_i$ are used to determine the nature of these quotients, just as was done for quotients (17), and it is found that

$$(P_i(x,\rho) - Q_i(x,\rho)) = \rho^{-\zeta-r-1} \sum_{j=0}^{\infty} g_{ij}(x)\rho^{-j} \qquad (i = 1, \cdots, n).$$

A sufficient number of hypotheses have been made to assure continuity in the functions under consideration; hence it is possible to select the positive constants S and R_2 sufficiently large so that

$$(24) \qquad \begin{aligned} &| P_i(x,\rho) - Q_i(x,\rho) | < S/\rho^{\zeta+r+1} \\ &\qquad (i = 1, \cdots, n;\ |\rho| > R_2 > R_1;\ a \leq x \leq b). \end{aligned}$$

Asymptotic approximations to solutions. To determine the way in which the functions $z_i(x,\rho)$ approximate solutions of (1), the original differential equation (1) is first rewritten

$$\mathfrak{M}(y) = \mathfrak{N}(y) \quad \text{with} \quad \mathfrak{N}(y) \equiv \mathfrak{M}(y) - \mathfrak{L}(y)$$

and this formally non-homogeneous differential equation is then replaced by the equivalent integral equation [6]

$$(25) \quad y(x,\rho) = \sum_{i=1}^{n} c_i z_i(x,\rho) + \int_a^x \Big(\sum_{i=1}^{n} z_i(x,\rho) Z_i(\xi,\rho) \mathfrak{N}(y(\xi,\rho)) \Big) d\xi.$$

Here the c_i's are arbitrary constants and the adjoint functions $Z_i(x,\rho)$ are defined by the system of equations

$$\sum_{i=1}^{n} z_i^{(j)}(x,\rho) Z_i(x,\rho) = \begin{cases} 0 & \text{for } j = 0, 1, \cdots, n-2 \\ 1 & \text{for } j = n-1. \end{cases}$$

It is readily deduced, after applying Cramer's rule, that the adjoint functions have the structure

$$Z_i(x,\rho) = \rho^{-r_i} H_i(x,\rho) \exp[-\int_a^x \Omega_i dt] \qquad (i = 1, \cdots, n).$$

[6] Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. 1, p. 78.

The constants τ_i are the same as those appearing in identities (23) and the symbols $H_i(x, \rho)$ designate uniformly bounded functions.

A positive constant H may be selected sufficiently large so that

$$|H_i(x, \rho)| < H \qquad (a \leq x \leq b;\ |\rho| > R_2),$$

and also

$$|\mathcal{D}_i{}^j \eta_i(x, \rho)| < H\,|\rho^{ir}| \qquad (i = 1, \cdots, n;\ j = 0, 1, \cdots, n-1).$$

Hence

$$|z_i{}^{(j)}(x, \rho) Z_i(\xi, \rho)| < H^2\,|\rho^{jr-\tau_i} \exp[\int_\xi^x \Omega_i dt]| \tag{26}$$
$$(i = 1, \cdots, n;\ j = 0, 1, \cdots, n-1).$$

A region defined by the inequalities

$$a \leq x \leq b,\ \ \gamma \leq \arg \rho \leq \theta,\ \text{ and }\ |\rho| > R_3 > R_2$$

will be called a Ξ-region, if it is possible to rearrange the subscripts on the Ω_i's in such a way that the real portions of the functions $\Omega_i(x, \rho)$ can be ordered as follows:

$$\Re(\Omega_i(x, \rho)) < \Re(\Omega_{i+1}(x, \rho))$$
$$\text{unless } \Omega_i(x, \rho) \equiv \Omega_{i+1}(x, \rho), \text{ and then } \Re(\Omega_i) = \Re(\Omega_{i+1})$$
$$(i = 1, 2, \cdots, n-1).$$

Within such a region

$$|\exp[\int_\xi^x \Omega_1 dt]| \leq |\exp[\int_\xi^x \Omega_2 dt]| \leq \cdots \leq |\exp[\int_\xi^x \Omega_i dt]| \tag{27}$$
$$\leq \cdots \leq |\exp[\int_\xi^x \Omega_n dt]|$$

for $x \geq \xi$.

The integral equation (25) and inequalities (24), (26), and (27) all find their counterpart in a paper by Birkhoff.[7] If we follow, therefore, Birkhoff's line of reasoning, we conclude that the integral equations

$$y_i(x, \rho) = z_i(x, \rho) + \int_a^x (\sum_{h=1}^{i} z_h(x, \rho) Z_h(\xi, \rho))\, \mathfrak{N}(y_i(\xi, \rho))\, d\xi$$
$$+ \int_b^x (\sum_{h=i+1}^{n} z_h(x, \rho) Z_h(\xi, \rho))\, \mathfrak{N}(y_i(\xi, \rho))\, d\xi$$
$$(i = 1, \cdots, n)$$

define a fundamental set of solutions for the original differential equation (1) with the property that within a Ξ-region

[7] G. D. Birkhoff, *op. cit.*, formulas (32), (29), (43), and (7) respectively.

$$(28) \qquad y_i^{(j)}(x,\rho) = z_i^{(j)}(x,\rho) + \frac{\mathcal{E}(x,\rho)\,\exp[\int_a^x \Omega_i(t,\rho)\,dt]}{\rho^{\zeta+\tau+1+\tau-r(n+j)}}$$

$$(a \leq x \leq b\,;\ |\rho| > R_3\,;\ i = 1,\cdots,n\,;\ j = 0,1,\cdots,n-1\,;\ \zeta > 2rn - 2r - 1).$$

The constant τ is the smallest of the set $\tau_1, \tau_2, \cdots, \tau_n$. It is relation (28) that is to replace (3) when multiple roots of the characteristic equation (4) are admitted. The functions $z_i(x,\rho)$ are therefore asymptotic approximations to the solutions $y_i(x,\rho)$.

The solutions indicated in (28) are dependent upon the integer ζ. This integer is fixed, it will be recalled, by the point at which the formal series are cut short. Tamarkin and Besikowitsch [8] have shown how a solution y_n can be chosen to satisfy a relation of type (28) independently of the value of ζ. This solution is then used to reduce the order of the given differential equation to $n-1$ and through mathematical induction it is found,[9] provided the characteristic roots are unequal, that a set of n solutions exist satisfying (28) independently of ζ. This line of reasoning can readily be adopted to the present situation in which multiple roots are admitted. It is possible, for instance, that $\Omega_{n-m+1}, \Omega_{n-m+2}, \cdots$, and Ω_n $(1 \leq m \leq n)$ are all identical. In this event m solutions $y_{n-m+1}, y_{n-m+2}, \cdots, y_n$ would exist which satisfy (28) independently of ζ. The induction would proceed, therefore, even more rapidly than in the classical case, for the m solutions $y_{n-m+1}, \cdots, y_n$ would be used to reduce the order of the differential equation (1) to $n-m$. Hence it is concluded that there exist n independent solutions of the original differential equation (1) which have within a Ξ-region the asymptotic expansions

$$y_i^{(j)} \sim \mathcal{D}_i^{\,j}(z_{i,0}(x) + z_{i,1}(x)/\rho^{1/\beta_i} + \cdots \text{to inf.})\,\exp[\int_a^x \Omega_i(t,\rho)\,dt]$$
$$(i = 1,\cdots,n\,;\ j = 0,1,\cdots,n-1)$$

where $\mathcal{D}_i^{\,j}$ represents, as before, the symbolic operator $(d_i/dx + \Omega_i(x,\rho))$ here applied j times in succession; and where the integers β_i, and the functions Ω_i and $z_{i,j}$ are the same as those appearing in the formal solutions.

UNIVERSITY OF WISCONSIN.

[8] J. Tamarkin and A. Besikowitsch, *Mathematische Zeitschrift*, vol. 21 (1924), pp. 119-123.

[9] O. Perron, *op. cit.*, pp. 14-21.

BIHARMONIC FUNCTIONS AND CERTAIN GENERALIZATIONS.[1]

By EDWARD KASNER.

In the following paper we study certain new characterizations of biharmonic functions in terms of conformal transformations; and certain generalizations connected with various subgroups of the total conformal group, and also with various subsets which are not groups. For example we deal not only with the six parameter inversion group, but also with the set of ∞^3 pure inversions. The main results are given in Theorems 1-8. I wish to acknowledge the valuable assistance of George Comenetz in writing this paper, especially in the proof of Theorem 5.

THE POINCARÉ EQUATIONS.

1. Poincaré applied the term biharmonic to those functions of four real variables $F(x, y, x_1, y_1)$ which can be regarded as the real part of an analytic function of two complex variables $z = x + iy$, $z_1 = x_1 + iy_1$. The same word biharmonic is also in use to denote those functions of two real variables $f(x, y)$ which satisfy the double Laplace equation $\Delta\Delta f = 0$, of the fourth order (connected with the problem of the bending of a flat plate);[2] but the two meanings have nothing to do with each other. Biharmonic is used in this paper exclusively in the sense of Poincaré.

If $F(x, y, x_1, y_1)$ is biharmonic, then a conjugate function $G(x, y, x_1, y_1)$ exists such that $F + iG$ is an analytic function of the complex variables $x + iy$ and $x_1 + iy_1$. Hence F and G satisfy the Cauchy-Riemann differential equations with respect to each of the two complex variables; thus we have four equations of the first order:

$$(1) \qquad F_x = G_y, \quad F_y = -G_x; \qquad F_{x_1} = G_{y_1}, \quad F_{y_1} = -G_{x_1}.$$

If we differentiate these relations partially with respect to x, y, x_1 and y_1, and eliminate G, we find that F must satisfy a system of four partial differential equations of second order:

$$(2) \qquad \begin{aligned} P_1F &\equiv F_{xx} + F_{yy} = 0, & P_2F &\equiv F_{xx_1} + F_{yy_1} = 0, \\ P_3F &\equiv F_{xy_1} - F_{yx_1} = 0, & P_4F &\equiv F_{x_1x_1} + F_{y_1y_1} = 0. \end{aligned}$$

(The symbols P_1, P_2, P_3, P_4 can be thought of as differential operators, which

[1] Abstracts in *Bulletin of the American Mathematical Society*, vol. 35 (1929), pp. 290, 444.

[2] Love, *Elasticity*, 4th ed., p. 135.

we may call the Poincaré operators. In fact, P_1 is identical with the Laplace operator Δ, taken with respect to the variables x and y, and P_4 is Δ with respect to x_1, y_1.) Since

$$(3) \qquad P_1F + P_4F \equiv \partial^2F/\partial x^2 + \partial^2F/\partial y^2 + \partial^2F/\partial x_1^2 + \partial^2F/\partial y_1^2 = 0,$$

we see that a biharmonic function is necessarily harmonic in the four variables x, y, x_1, y_1; but the converse is not true: a harmonic function of x, y, x_1, y_1 need not be the real part of any analytic function of $x + iy$ and $x_1 + iy_1$ (in contrast to the case of one complex variable), for equations (2) do not follow from (3). Equations (2) are sufficient as well as necessary for F to be biharmonic.[3] They were first given by Poincaré.[4]

The Conformal Group.

2. If in an analytic function of two complex variables z and z_1, we replace z_1 by an analytic function of z, the result is an analytic function of z alone. If we consider the real part of the function of z and z_1, we see that this remark implies the following theorem on biharmonic functions:

Let

$$(4) \qquad T: \quad x_1 = \phi(x, y), \quad y_1 = \psi(x, y)$$

be a substitution replacing x_1 and y_1 by functions of x and y. Then $F(x, y, x_1, y_1)$ reduces to a function K depending only on x and y:

$$(5) \qquad F(x, y, x_1, y_1) = F[x, y, \phi(x, y), \psi(x, y)] \equiv K(x, y).$$

Now replacing z_1 by an analytic function of z corresponds to using a *conformal* substitution T. Hence the obvious theorem we obtain is this:

Theorem 1. *If a conformal substitution $x_1 = \phi(x, y)$, $y_1 = \psi(x, y)$ is performed on a biharmonic function $F(x, y, x_1, y_1)$, then F becomes a harmonic function $K(x, y)$.*

We now give two fundamental converses of this theorem:

Theorem 2. *Biharmonic functions are the only functions of four real variables which are converted into harmonic functions by* every *conformal substitution.*

---

[3] Osgood, *Funktionentheorie* II, 1, p. 22. (The region of definition of F must satisfy a certain connectivity restriction.)

[4] *Comptes Rendus*, vol. 96 (1883), p. 238; *Acta Mathematica*, vol. 2 (1883), p. 99; vol. 22 (1898), p. 112.

THEOREM 3. *The only substitutions which convert* ALL *biharmonic functions into harmonic functions are the conformal substitutions.*

To prove Theorems 2 and 3 we need the expression for $K_{xx} + K_{yy}$, or ΔK, in terms of F, ϕ, and ψ. This is found by differentiating (5) partially with respect to x and y (for example, $K_x = F_x + F_{x_1}\phi_x + F_{y_1}\psi_x$, etc.). The result is found to be

$$(6)\quad \Delta K = F_{xx} + F_{yy} + 2F_{xx_1}\phi_x + 2F_{xy_1}\psi_x + 2F_{yx_1}\phi_y + 2F_{yy_1}\psi_y + F_{x_1x_1}(\phi_x^2 + \phi_y^2) + 2F_{x_1y_1}(\phi_x\psi_x + \phi_y\psi_y) + F_{y_1y_1}(\psi_x^2 + \psi_y^2) + F_{x_1}\Delta\phi + F_{y_1}\Delta\psi.$$

In the case of a conformal substitution T, the conformality conditions $\phi_x = \psi_y$, $\phi_y = -\psi_x$ can be used to simplify (6) by the elimination of the derivatives. of ψ. When T is conformal, (6) in fact reduces to *our principal formula*

$$(7)\quad \Delta K = P_1F + 2\phi_xP_2F - 2\phi_yP_3F + (\phi_x^2 + \phi_y^2)P_4F.$$

Now in Theorem 2 we suppose that F is a function such that K becomes harmonic no matter what conformal substitution T is used. That is,

$$(8)\quad [P_1 + 2\phi_xP_2 - 2\phi_yP_3 + (\phi_x^2 + \phi_y^2)P_4]F(x, y, \phi, \psi) = 0$$

identically in x and y, for every conformal ϕ, ψ. We must show that F is biharmonic.

Since (8) holds for any conformal T, we may use, in particular, linear conformal substitutions, or similitudes:

$$(9)\quad x_1 = ax - by + c, \qquad y_1 = bx + ay + d,$$

(where a, b, c, d are arbitrary constants). Then (8) becomes an identity in the six independent variables x, y, a, b, c, d. In place of c and d, we may introduce x_1 and y_1 as independent variables by solving equations (9) for c, d in terms of x, y, a, b, x_1, y_1. Then (8) becomes

$$(10)\quad [P_1 + 2aP_2 + 2bP_3 + (a^2 + b^2)P_4]F(x, y, x_1, y_1) = 0,$$

an identity in x, y, x_1, y_1, a, b. Varying a and b, we see that $P_1F, \cdots, P_4F$ must vanish at each fixed point x, y, x_1, y_1 in the region of definition of F. Hence F must be biharmonic. This proves Theorem 2.

For Theorem 3, we suppose that T is a substitution which turns every biharmonic F into a harmonic K. We must show that T is conformal.

Since the right member of (6) vanishes for any biharmonic F, it vanishes in particular when F is linear in x_1 and y_1 with constant coefficients (any linear

function of x, y, x_1, y_1 is of course biharmonic). In that case, all the terms of (6) drop out except the last two; and the coefficients of these two terms, F_{x_1} and F_{y_1}, reduce to arbitrary constants. This shows that $\Delta\phi = \Delta\psi = 0$. We next employ the particular biharmonic functions $F = a(xx_1 - yy_1) - b(xy_1 + yx_1)$, where a and b are arbitrary constants (the real part of $(a + bi)zz_1$). Substituting in (6), we have

$$2a(\phi_x - \psi_y) - 2b(\phi_y + \psi_x) = 0. \tag{11}$$

Hence $\phi_x - \psi_y = \phi_y + \psi_x = 0$, so that T must be conformal. This proves Theorem 3.

Subgroups and Subsets.

3. A function $F(x, y, x_1, y_1)$ which is made harmonic by every conformal substitution T is necessarily biharmonic, by Theorem 2. The proof of this theorem shows, in fact, that F must actually be biharmonic even if we merely require F to be converted into a harmonic function by the subgroup of the conformal group consisting of all similitudes.

Theorem 2′. *The only functions of four real variables which are converted into harmonic functions by every similitude are the biharmonic functions.*

If other *subgroups* or even *subsets which are not groups*, of the conformal group are used, however, we may expect that other functions of four variables, besides the biharmonic, will satisfy the requirement of becoming harmonic under every substitution of the subset. Larger classes of functions will thus be obtained, which we may then regard as generalizations of the biharmonic class.

We consider this problem in the case of the following subsets. (For completeness, we include similitudes in the list. The equations of substitution are given both in the complex form and in the equivalent real form (4). The letters a, b, c, d, θ stand for arbitrary real constants; and $\lambda = a + ib$, and $\mu = c + id$.)

(I) The group of ∞^4 similitudes: $z_1 = \lambda z + \mu$, or equations (9).

(II) The group of ∞^3 magnifications: $z_1 = az + \mu$, or

$$x_1 = ax + c, \qquad y_1 = ay + d. \tag{12}$$

(III) The group of ∞^3 rigid motions: $z_1 = e^{i\theta}z + \mu$, or

$$x_1 = x\cos\theta - y\sin\theta + c, \qquad y_1 = x\sin\theta + y\cos\theta + d. \tag{13}$$

(IV) The group of ∞^2 translations: $z_1 = z + \mu$, or

(14) $$x_1 = x + c, \qquad y_1 = y + d.$$

(V) The set of ∞^2 central symmetries: $z_1 = -z + \mu$, or

(15) $$x_1 = -x + c, \qquad y_1 = -y + d.$$

(VI) The group of $2\infty^2$ translations and central symmetries: $z_1 = \pm z + \mu$, or

(16) $$x_1 = \pm x + c, \qquad y_1 = \pm y + d.$$

We state the answers obtained in these six cases in a single theorem (the symbol *bih* stands for an arbitrary biharmonic function, the symbol $\mathfrak{R}$ for the real part of an analytic function of complex variables, c for an arbitrary constant; $\bar{z} = x - iy$ and $\bar{z}_1 = x_1 - iy_1$).

THEOREM 4. *The class of functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by all the substitutions $x_1 = \phi(x, y)$, $y_1 = \psi(x, y)$ of a subset of the conformal group is, in the cases of the subsets* I *to* VI, *the following:*

(I) (*Similitudes*): $F = bih(x, y, x_1, y_1)$.

(II) (*Magnifications*): $F = bih(x, y, x_1, y_1) + c(xy_1 - yx_1)$.

(III) (*Motions*): $F = bih(x, y, x_1, y_1) + c(x^2 + y^2 - x_1^2 - y_1^2)$.

(IV) (*Translations*): $F = g(x + x_1, x - x_1, y + y_1, y - y_1)$, *where g is harmonic in the pair of variables $x + x_1$, $y + y_1$, and arbitrary in $x - x_1$, $y - y_1$; or* $F = \mathfrak{R}[f(z, z_1, \bar{z} - \bar{z}_1)]$.

(V) (*Central symmetries*): $F = g(x + x_1, x - x_1, y + y_1, y - y_1)$, *where g is harmonic in $x - x_1$, $y - y_1$, and arbitrary in $x + x_1$, $y + y_1$; or* $F = \mathfrak{R}[f(z, z_1, \bar{z} + \bar{z}_1)]$.

(VI) (*Translations and central symmetries*):

$$F = g(x + x_1, x - x_1, y + y_1, y - y_1),$$

where g is harmonic in the pairs $x + x_1$, $y + y_1$, and $x - x_1, y - y_1$; or

$$F = \mathfrak{R}[f_1(z, z_1) + f_2(z + z_1, \bar{z} - \bar{z}_1)];$$

or

$$F = bih_1(x, y, x_1, y_1) + bih_2(x, y_1, x_1, y).^5$$

---

[5] Two other cases may be mentioned: the group of ∞^2 similitudes leaving the origin fixed, $z_1 = \lambda z$ which gives $F = \mathfrak{R}[f(z, z_1, \bar{z}/\bar{z}_1)]$; and the set of ∞^2 reciprocations $z_1 = \lambda/z$ which gives $F = \mathfrak{R}[f(z, z_1, z\bar{z}_1)]$.

Proof of Theorem 4.

Case I has already been settled in Theorem 2′.

To prove Theorem 4 in cases II to VI we begin as in the proof of Theorem 2 (i. e., of case I). Equation (8) must hold identically in x and y, for every ϕ, ψ of types (12) to (16) respectively. Substituting (12)-(16) into (8), we obtain identities in x, y and some of the letters a, b, c, d, θ. If we solve (12)-(16) for c and d in terms of x_1, y_1, and the other letters, we can eliminate c and d from these identities, introducing x_1 and y_1 as independent variables instead. Now if we think of x, y, x_1, y_1 as fixed, and vary a, b, θ, we find that the conditions which F must satisfy are, in the various cases, as follows:

$$\begin{aligned}
&\text{(II)} && P_1F = P_2F = P_4F = 0.\\
&\text{(III)} && (P_1 + P_4)F = P_2F = P_3F = 0.\\
&\text{(IV)} && (P_1 + 2P_2 + P_4)F = 0.\\
&\text{(V)} && (P_1 - 2P_2 + P_4)F = 0.\\
&\text{(VI)} && (P_1 + P_4)F = P_2F = 0.
\end{aligned}$$

These can be looked on as systems of partial differential equations of second order, to be solved for the unknown F. Fortunately all these equations are linear, with constant coefficients, and can be solved explicitly.

The solution, in cases II and III, can be carried out most conveniently with the aid of the general identities

$$\text{(17)} \qquad \begin{aligned} (P_1)_{x_1} &= (P_2)_x - (P_3)_y, & (P_4)_x &= (P_2)_{x_1} + (P_3)_{y_1},\\ (P_1)_{y_1} &= (P_2)_y + (P_3)_x, & (P_4)_y &= (P_2)_{y_1} - (P_3)_{x_1}. \end{aligned}$$

These are easy to verify from the definitions of $P_1, \cdots, P_4$ in (2).

First, under the conditions of case II, we see from (17) that P_3F is independent of x, y, x_1 and y_1. Hence $P_1F = P_2F = P_4F = 0$, $P_3F = 2c$ (an arbitrary constant). A particular solution of this non-homogeneous system is: $F = c(xy_1 - yx_1)$. The general solution of the corresponding homogeneous system $P_1F = P_2F = P_3F = P_4F = 0$ is: $F =$ *any biharmonic function.* Hence the general solution of the equations of case II is:

$$F = bih(x, y, x_1, y_1) + c(xy_1 - yx_1),$$

which is the result given in Theorem 4.

In case III we find from (17) that P_1F depends only on x and y, while P_4F depends only on x_1 and y_1. Since $P_1F + P_4F = 0$, in view of the different variables in P_1F and P_4F we must have $P_1F = -P_4F = 4c$. The general solution of the non-homogeneous system of equations is now obtained just as in case II.

The equations of cases IV, V, VI are simplified by the introduction of the new variables

$$\begin{aligned} \xi &= x + x_1, & \eta &= y + y_1, \\ \xi_1 &= x - x_1, & \eta_1 &= y - y_1. \end{aligned} \tag{18}$$

Transforming the equations to these new variables, we find for IV: $F_{\xi\xi} + F_{\eta\eta} = 0$, and for V: $F_{\xi_1\xi_1} + F_{\eta_1\eta_1} = 0$. Hence in case IV, F is harmonic in ξ and η, and in case V, F is harmonic in ξ_1 and η_1, in accordance with Theorem 4.

The first statement under case VI in Theorem 4 follows from the results in cases IV and V. Proofs of the remaining statements in Theorem 4 are omitted. We remark that the work is simplified by the use of the minimal coördinates z, $\bar{z}$, z_1, $\bar{z}_1$ instead of x, y, x_1, y_1, F then being assumed analytic (see the proof of Theorem 5 below). In the proofs above it was sufficient to assume the existence of continuous third derivatives of F. (Actually all the results are valid without assuming analyticity.)

ANTI-BIHARMONIC FUNCTIONS AND REVERSE CONFORMAL TRANSFORMATIONS.

4. The real part of an analytic function of z and z_1 is, by definition, biharmonic; and the real part of an analytic function of the conjugate complex variables $\bar{z}$ and $\bar{z}_1$ is also biharmonic, since in the latter case equations (1) hold with G replaced by $-G$, and then equations (2) still follow. But the real part of an analytic function of z and $\bar{z}_1$, or of $\bar{z}$ and z_1, is in general not biharmonic; we shall term it "anti-biharmonic." The sign of G is changed in only one pair of equations (1), and then instead of (2) we find

$$\begin{aligned} P_1F &\equiv F_{xx} + F_{yy} = 0, & P'_2F &\equiv F_{xx_1} - F_{yy_1} = 0, \\ P'_3F &\equiv F_{xy_1} + F_{yx_1} = 0, & P_4F &\equiv F_{x_1x_1} + F_{y_1y_1} = 0 \end{aligned} \tag{19}$$

as the *defining equations of anti-biharmonic functions.*

A "reverse conformal" substitution T is one obtained by setting z_1 equal to an analytic function of $\bar{z}$. It is easy to modify the proofs of Theorems 1, 2, and 3 so as to show that the same theorems still hold with "anti-biharmonic" and "reverse conformal" in place of "biharmonic" and "conformal," respectively.

The question considered in Theorem 4 can also be raised here: given a set of reverse conformal substitutions, to determine the functions of x, y, x_1, y_1 which are turned into harmonic functions of x, y by every substitution of the set. Theorem 5 answers this question for the following important sets of substitutions:

(VII) The total set of ∞^2 line symmetries: $z_1 = -\bar{z}\, e^{2i\theta} + ae^{i\theta}$, or

$$\begin{aligned} x_1 &= -x\cos 2\theta - y\sin 2\theta + a\cos\theta, \\ y_1 &= -x\sin 2\theta + y\cos 2\theta + a\sin\theta. \end{aligned} \tag{20}$$

(VIII) The total set of ∞^3 inversions: $z_1 = \dfrac{r^2}{\bar{z} - \bar{\lambda}} + \lambda$, or

$$\begin{aligned} x_1 &= \frac{r^2(x-a)}{(x-a)^2 + (y-b)^2} + a, \\ y_1 &= \frac{r^2(y-b)}{(x-a)^2 + (y-b)^2} + b. \end{aligned} \tag{21}$$

(Equations (20) represent a reflection in the line whose normal makes an angle θ with the x-axis, and which is at the distance $a/2$ from the origin; (21) are the equations of an inversion in the circle of radius r, with center at the point $\lambda \equiv a + ib$. This can be seen directly from the complex forms of (20) and (21)).

Symmetries and Inversions.

Theorem 5. *The class of functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every one of a set of reverse conformal substitutions is, in the case of sets* VII *and* VIII, *the following:*

(VII) (*Line symmetries*) : $F = \mathfrak{R}\left[f\left(z, \bar{z}_1, \dfrac{z - z_1}{\bar{z} - \bar{z}_1}\right)\right]$.

(VIII) (*Inversions*) : $F = \mathfrak{R}\left[f(z, \bar{z}_1) + \text{const.} \cdot \log \dfrac{z - z_1}{\bar{z} - \bar{z}_1}\right]$,

that is, $F = antibih(x, y, x_1, y_1) + c\tan^{-1}\dfrac{y - y_1}{x - x_1}$.[6]

These two problems seem important on account of their possible connection with automorphic functions. It will be noticed that in both cases we are dealing with subsets which do not have the group property. The analytic work is complicated since it turns out that we must solve linear partial differential equations with variable coefficients.

To prove Theorem 5 we first derive the condition, analogous to (8), for a given function $F(x, y, x_1, y_1)$ to be reduced to a harmonic function of x and y by a *reverse* conformal substitution. This is found from (6) by using the reverse conformality conditions, $\phi_x = -\psi_y$, $\phi_y = \psi_x$. The result is

$$P_1F + 2\phi_x P'_2F + 2\phi_y P'_3F + (\phi_x^2 + \phi_y^2)P_4F = 0. \tag{22}$$

[6] The functions which are made harmonic by line symmetries include those made harmonic by inversions. This is because line symmetries are the limiting case of inversions, when the circle of inversion becomes a straight line.

Now following the procedure explained in the proof of Theorem 2, we substitute into (22) the special ϕ's of (20) and (21), and then eliminate a, b, θ by solving (20) and (21) for these parameters. In case VII we find the single equation

$$(23)\quad [(x-x_1)^2+(y-y_1)^2](P_1+P_4)F-2[(x-x_1)^2-(y-y_1)^2]P'_2F \\ -4(x-x_1)(y-y_1)P'_3F=0.$$

In case VIII, by varying the remaining parameter r we arrive at the system of three equations

$$(24)\quad P_1F=P_4F=[(x-x_1)^2-(y-y_1)^2]P'_2F \\ +2(x-x_1)(y-y_1)P'_3F=0.$$

As before, these are all linear partial differential equations of second order in the unknown $F(x, y, x_1, y_1)$, but the coefficients are no longer constants. However, it is still possible to integrate the equations explicitly. The first step is to simplify them by introducing the minimal variables

$$(25)\quad u=x+iy,\quad v=x-iy,\quad u_1=x_1+iy_1,\quad v_1=x_1-iy_1\ ^{7}$$

(the notation u, v, u_1, v_1 is more convenient than $z, \bar{z}, z_1, \bar{z}_1$). When this change of variables is carried out ($P_1F\equiv F_{xx}+F_{yy}=4F_{uv}$, etc.), the equations assume the forms

$$(26)\quad (\text{VII})\quad (u-u_1)^2F_{uu_1}-(u-u_1)(v-v_1)(F_{uv}+F_{u_1v_1})+(v-v_1)^2F_{vv_1}=0,$$

$$(27)\quad (\text{VIII})\quad F_{uv}=F_{u_1v_1}=(u-u_1)^2F_{uu_1}+(v-v_1)^2F_{vv_1}=0.$$

Now in case VII we find that equation (26) can be factored symbolically as follows:

$$(28)\quad \left[(u-u_1)\frac{\partial}{\partial u}-(v-v_1)\frac{\partial}{\partial v_1}-1\right]\left[(u-u_1)\frac{\partial}{\partial u_1}-(v-v_1)\frac{\partial}{\partial v}\right]F=0.$$

Hence if we denote by E the function obtained by applying to F the operator in the second pair of brackets, we may replace (28) by the pair of linear equations

$$(29)\quad (u-u_1)F_{u_1}-(v-v_1)F_v=E,$$

$$(30)\quad (u-u_1)E_u-(v-v_1)E_{v_1}-E=0.$$

The second of these is easily solved, with the result

[7] I. e., F is assumed analytic in the four real variables x, y, x_1, y_1, and this analytic function is extended to complex values of x, y, x_1, y_1. Thus F stands for a function of complex variables from this point up to the last step of the proof, when x, y, x_1, y_1 are specialized to real values again.

$$E = (u - u_1) f_1\left(v, u_1, \frac{v - v_1}{u - u_1}\right), \tag{31}$$

where f_1 is an arbitrary function of three variables. Substituting (31) into (29) and solving this in turn (for instance, with the aid of $\frac{v - v_1}{u - u_1}$ as an independent variable in place of u), we find

$$F = f_2\left(u, v_1, \frac{u - u_1}{v - v_1}\right) + f_3\left(v, u_1, \frac{v - v_1}{u - u_1}\right). \tag{32}$$

Here f_2 and f_3 are arbitrary, except of course that their sum must be real when we now restrict x, y, x_1, y_1 to real values. Since the three variables appearing in f_3 now become the conjugates of those in f_2, an equivalent form of (32) is

$$F = \Re\left[f\left(u, v_1, \frac{u - u_1}{v - v_1}\right)\right], \tag{33}$$

where f is arbitrary (this follows from the fact that the real part of an analytic function of complex variables is also the real part of some analytic function of the conjugates of those variables). The result given in Theorem 5 is identical with (33).

In case VIII, $F_{uv} = F_{u_1 v_1} = 0$ by (27). This means that

$$F = f_1(u, u_1) + f_2(u, v_1) + f_3(v, u_1) + f_4(v, v_1), \tag{34}$$

where the f's are arbitrary. Substituting (34) into the third equation in (27) and noting the separation of variables, we have

$$(u - u_1)^2 (f_1)_{uu_1} = -(v - v_1)^2 (f_4)_{vv_1} = \lambda, \tag{35}$$

where λ is a complex constant. Integrating (35) and substituting in (34), we find

$$F = f_4(u, v_1) + f_5(v, u_1) + \lambda \log \frac{u - u_1}{v - v_1}. \tag{36}$$

The functions f_4 and f_5 and the constant λ in (36) are arbitrary, subject to the condition that when we now make x, y, x_1, y_1 real, F also becomes real. Since the variables in f_5 are now the conjugates of those in f_4, and since the coefficient of λ is now a pure imaginary, it follows that a form of the answer equivalent to (36) is

$$F = \Re[f(u, v_1)] + \Re\left[\lambda \log \frac{u - u_1}{v - v_1}\right], \tag{37}$$

where the function f and the constant λ are completely arbitrary. The results stated in case VIII of Theorem 5 are readily derived from (37). This concludes the proof of Theorem 5.

Mixed Groups.

5. The previous theorems deal either with groups and sets of conformal substitutions, or else with sets of reverse conformal substitutions. There remains the case of mixed groups or sets, consisting of both direct and reverse conformal substitutions. In particular, we may ask: what functions are rendered harmonic by all the substitutions of the total (direct and reverse) conformal group? By Theorem 2 and its analogue for reverse conformal substitutions, the required functions are those which are at once biharmonic and anti-biharmonic; that is, they are those functions which satisfy equations (2) and also (19). Hence we must have

$$F_{xx} + F_{yy} = F_{xx_1} = F_{xy_1} = F_{yx_1} = F_{yy_1} = F_{x_1x_1} + F_{y_1y_1} = 0. \tag{38}$$

The solution of (38) is

$$F = f(x, y) + g(x_1, y_1), \tag{39}$$

where f and g are arbitrary harmonic functions of two variables. This proves

Theorem 6. *The only functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every substitution of the total (direct and reverse) conformal group, are those which are equal to the sum of a harmonic function of x and y and a harmonic function of x_1 and y_1.*

Finally, we take up the question of determining what functions are made harmonic by all the substitutions of a subgroup of the total conformal group, in the case of these two mixed subgroups:

(IX) The total group of $2\infty^3$ rigid motions, generated by the set VII of line symmetries (the group of ordinary plane geometry).

(X) The inversion group, generated by the set VIII of transformations by reciprocal radii vectores. This consists of the $2\infty^6$ linear fractional transformations of z and linear fractional transformations of $\bar{z}$.

A function F which is made harmonic by the total group of motions IX must satisfy the equations deduced in case III for direct motions, together with the corresponding equations for reflected motions. The latter equations are derived from the former by replacing P_2, P_3 by P'_2, P'_3, as we see by comparing (22) with (8). Hence F is determined by the system

$$F_{xx} + F_{yy} + F_{x_1x_1} + F_{y_1y_1} = F_{xx_1} = F_{xy_1} = F_{yx_1} = F_{yy_1} = 0. \tag{40}$$

The last four equations mean that F is separable in the pairs x, y and x_1, y_1. Then using this separation of variables in the first equation, we see as in the proof of III that F equals the sum of a harmonic function of x and y,

a harmonic function of x_1 and y_1, and a function of the special form $c(x^2+y^2-x_1^2-y_1^2)$.

In case X, it is sufficient to recall that the inversion group contains all similitudes, both direct and reverse. Then by Theorem 2′ and its analogue for reverse similitudes, F must be both biharmonic and anti-biharmonic. Hence the answer for the inversion group is the same as for the total conformal group. We have now proved

THEOREM 7. *The class of functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every substitution $x_1=\phi(x,y)$, $y_1=\psi(x,y)$ of a mixed group of (direct and reverse) conformal substitutions is, in the case of the groups* IX *and* X, *the following:*

(IX) (*Total group of motions*):
$F=f(x,y)+g(x_1,y_1)+c(x^2+y^2-x_1^2-y_1^2)$, *where f and g are arbitrary harmonic functions and c is an arbitrary constant.*

(X) (*The inversion group*): $F=f(x,y)+g(x_1,y_1)$, *where f and g are arbitrary harmonic functions.*

Since the group of motions IX is generated by the set of line symmetries VII, and therefore includes this set, the class of functions F which are made harmonic by the group IX must be included in the class of functions made harmonic by the set VII. The same thing may be said for the inversion group X and the set of inversions VIII. That is, the answers in Theorem 5 should include, respectively, the answers in Theorem 7. For VIII and X this is evidently so, for the answer to case X can also be written in the form

$$F=\mathfrak{R}[f_1(z)+f_2(\bar{z}_1)], \tag{41}$$

and this is of the type VIII in Theorem 5. For cases VII and IX the fact is less obvious, but it follows from the relation

$$x^2+y^2-x_1^2-y_1^2=\mathfrak{R}\left[z^2\frac{\bar{z}-\bar{z}_1}{z-z_1}-\bar{z}_1^2\frac{z-z_1}{\bar{z}-\bar{z}_1}\right]. \tag{42}$$[8]

QUADRATIC BIHARMONIC FUNCTIONS.

6. By Theorem 2, not every (direct) conformal substitution can reduce a given function $F(x, y, x_1, y_1)$ to a harmonic function of x and y (unless of course F happens to be biharmonic). But for a *given* non-biharmonic F, there may still be some subset of the conformal group which makes F harmonic.

[8] This relation can be derived by following the steps of the proof of VII, taking F to be $x^2+y^2-x_1^2-y_1^2$. Of course it can also be verified directly.

We determine this subset in the case that F is a quadratic polynomial. There is no need for the subset to constitute a group, and we find that as a rule groups do not appear. We first establish the following

LEMMA. *If F is a quadratic polynomial in four variables, and if a conformal substitution T reduces F to a harmonic function, then T must be linear (unless F is biharmonic).*

For (8) must hold, with the P_iF reduced to constants since F is a quadratic polynomial. Differentiating (8) with respect to x and y, and rearranging terms, we have:

$$\begin{aligned} &\phi_{xx}(\phi_x P_4 + P_2)F + \phi_{xy}(\phi_y P_4 - P_3)F = 0, \\ &\phi_{xy}(\phi_x P_4 + P_2)F + \phi_{yy}(\phi_y P_4 - P_3)F = 0. \end{aligned} \tag{43}$$

Now since T is conformal, $\phi_{xx} + \phi_{yy} = 0$. If the determinant $\phi_{xx}\phi_{yy} - \phi_{xy}^2$ of (43) vanishes, then using $\phi_{yy} = -\phi_{xx}$, we have $\phi_{xx}^2 + \phi_{xy}^2 = 0$, so that $\phi_{xx} = \phi_{xy} = 0$. Then ϕ_{yy} is also zero, and T must be linear. If the determinant of the homogeneous system (43) does not vanish, then

$$\phi_x P_4 F + P_2 F = \phi_y P_4 F - P_3 F = 0.$$

This means that ϕ_x and ϕ_y are constants, so that ϕ, and therefore ψ, is linear; unless $P_4F = P_2F = P_3F = 0$. In the latter case we have from (8), $P_1F = 0$; hence F is biharmonic. This proves the lemma.

Now if a linear conformal substitution (9) makes F harmonic, the condition which the coefficients a, b in (9) must satisfy is, as in (10),

$$(a^2 + b^2)P_4F + 2aP_2F + 2bP_3F + P_1F = 0, \tag{44}$$

where $P_1F, \cdots, P_4F$ are independent constants, not all zero unless F is biharmonic. If we interpret a, b as the cartesian coördinates of a point in the plane, (44) is the general circle [9] equation; its (real) graph is either a circle, or a straight line, or a point, or does not exist in the real domain. Corresponding to each point (a, b) of the graph there are ∞^2 similitudes (9), since c, d are unrestricted. We can state this result in the following way:

THEOREM 8. *If F is a quadratic polynomial in x, y, x_1, y_1, the set of (direct) conformal substitutions which make F harmonic is one of four kinds, depending on the coefficients of F:*

---

[9] This may be called the *auxiliary circle* and we shall use it elsewhere. The proof of Theorem 2 can easily be changed to yield this stronger form of Theorem 2′: *If F is made harmonic by $4\infty^2$ similitudes situated "at four points" (a, b), and if these four points are not co-circular, then F is biharmonic.* For any given function $F(x, y, x_1, y_1)$ and any given bipoint (x, y), (x_1, y_1) we have a definite auxiliary circle.

(A) *The entire conformal group (namely, when F is biharmonic).*
(B) ∞^3 *similitudes "on a circle" or "on a straight line."*
(C) ∞^2 *similitudes "at a point."*
(D) *None.*

We remark that (B) and (D) constitute the general case; (C) is special, and (A) is still more special. For (C) arises only when the constants P_iF satisfy a certain relation, expressing the fact that the radius of the circle (44) is zero; and (A) requires the four conditions $P_iF = 0$.

The only groups under (B) are the group of rigid motions III ("on the circle" $a^2 + b^2 = 1$), and the group of magnifications II ("on the line" $b = 0$, excluding the origin); under (C) there is only the group of translations IV ("at the point" $(1, 0)$). To prove this we use the fact that if a_1, b_1 and a_2, b_2 are the coefficients of two similitudes (9), and if a_3, b_3 are the coefficients of the product of those two similitudes, then

$$a_3 + ib_3 = (a_1 + ib_1)(a_2 + ib_2).$$

This is evident from the complex form of (9), namely:

$$z_1 = (a + ib)z + (c + id).$$

Hence if the ∞^3 similitudes corresponding to points (a, b) on a circle are to form a group, the set of complex numbers $a + ib$ must form a group with respect to the operation of multiplication. Now it is easy to see that the successive powers of a complex number cannot lie on any circle, unless the modulus of the number is unity (or zero). Therefore the unit circle $a^2 + b^2 = 1$ is the only one possible; and this circle does in fact correspond to the group of motions. Similarly, the only straight line group is the real axis $b = 0$ (excluding 0); and the only point group (other than zero) is the point unity. These correspond, respectively, to the magnifications $z_1 = az + (c + id)$, and the translations $z_1 = z + (c + id)$.

Extensions.

7. We remark in conclusion that most of our theorems can be extended to *multi-harmonic* functions connected with functions of n (instead of two) complex variables. There is also a somewhat analogous theory of the reduction of polygenic (non-analytic) functions of complex variables to monogenic functions by monogenic substitutions.

Columbia University,
New York.

THE PROBLEM OF BOLZA IN THE CALCULUS OF VARIATIONS IN PARAMETRIC FORM.[1]

By MAGNUS R. HESTENES.[2]

1. Introduction. In a Riemannian space $\mathcal{R}$ as defined by Morse (X, pp. 107-108)[3] with local coördinates $(x) = (x^0, x^1, \cdots, x^n)$ the problem of Bolza can be defined as follows. Let $(x^1) = (x^{01}, x^{11}, \cdots, x^{n1})$ and $(x^2) = (x^{02}, x^{12}, \cdots, x^{n2})$ be respectively the end points 1 and 2 of the arc

$$x^i = x^i(t) \qquad (t^1 \leqq t \leqq t^2;\ i = 0, 1, \cdots, n). \tag{1.1}$$

We seek to find in a class of arcs (1.1) and sets of constants $(\alpha) = (\alpha^1, \cdots, \alpha^r)$ satisfying the differential equations and end conditions

$$\phi_\beta(x, \dot{x}) = 0 \qquad (\beta = 1, \cdots, m < n), \tag{1.2}$$

$$x^{is} = x^{is}(\alpha^1, \cdots, \alpha^r) \qquad (s = 1, 2) \tag{1.3}$$

one which minimizes a functional of the form

$$J = \theta(\alpha^1, \cdots, \alpha^r) + \int_{t^1}^{t^2} f(x, \dot{x})\, dt.$$

The problem of Bolza in non-parametric form has been studied by a number of writers.[4] The only treatments of the parametric problem known to the author are those of Eshleman (III) and Hefner (VII) in the problem of Lagrange.

Sufficient conditions for a minimum in the non-parametric problem without normality assumptions on the subintervals were first obtained by the author while he was a Research Assistant to Professor Bliss. These results were later extended (XIII) to the case in which no normality assumptions were made. This extension is a trivial one. More recently, Morse (XIV) and Reid (XV) obtained independently further sufficient conditions involving conjugate points which are equivalent to those given by the author (See XVI). In the present paper we give similar sufficiency conditions for a minimum in the parametric case.

---

[1] Presented to the American Mathematical Society, September 6, 1934.

[2] National Research Fellow 1933-4.

[3] Roman numerals in parentheses refer to the list of references at the end of this paper.

[4] For references other than those given at the end of this paper see (XIII).

2. Preliminary remarks. In the present paper we shall assume that two systems of coördinates (x) and (z) which admissibly represent the same neighborhood of a point of our Riemannian space $\mathfrak{R}$ are connected by an equation of the form

$$(2.1) \qquad z^i = z^i(x^0, x^1, \cdots, x^n) \qquad (i = 0, 1, \cdots, n)$$

in which the functions $z^i(x)$ are of class [5] C^4 and possess a non-vanishing jacobian. Let $(r) = (r^0, r^1, \cdots, r^n)$ be a contravariant vector, that is, suppose that under the transformation (2.1) the vector (r) is transformed into a vector (σ) according to the law [6]

$$(2.2) \qquad \sigma^i = \frac{\partial y^i}{\partial x^j} r^j \qquad (i, j = 0, 1, \cdots, n).$$

The functions $f(x, r)$, $\phi_\beta(x, r)$ appearing in equations (1.2) and (1.4) are assumed to be defined over $\mathfrak{R}$ for all sets $(r) \neq (0)$ and to be invariants under the transformations (2.1) and (2.2). Moreover these functions are assumed to be of class C^3 and to satisfy the homogeneity relations

$$(2.3) \qquad f(x, kr) = kf(x, r), \qquad \phi_\beta(x, kr) = k\phi_\beta(x, r) \qquad (k > 0).$$

By a *regular arc* is meant a set of points in $\mathfrak{R}$ defined locally by equations of the form (1.1) of class C^1 and having $(\dot{x}) \neq (0)$. By a *differentially admisssible arc* is meant a continuous succession of a finite number of regular arcs satisfying the differential equations (1.2). By an *admissible arc* is meant a differentiably admissible arc (1.1) and a set of constants $(\alpha) = (\alpha^1, \alpha^2, \cdots, \alpha^r)$ which satisfy the end conditions (1.3).

We are concerned with a particular admissible arc g without corners. As was noted by Morse we may assume without loss of generality that the arc g does not intersect itself. We suppose further that the matrix $\| \phi_{\beta r^i} \|$ has rank m along g and that the set (α) belonging to g is the set $(\alpha) = (0)$. The functions $x^{i_1}(\alpha)$, $x^{i_2}(\alpha)$, $\theta(\alpha)$ appearing in equations (1.3) and (1.4) are assumed to be of class C^2 near $(\alpha) = (0)$.

Our problem is then that of finding under what conditions the arc g will furnish a minimum to the functional J relative to neighboring admissible arcs. These conditions must be independent of the admissible coördinates (x) and of admissible parameters t and (α) used. Moreover, it can be shown that a neighborhood of g can be covered by a single coördinate system. In fact this

[5] A function $f(x^0, x^1, \cdots, x^n)$ is said to be of class C^n if it possesses continuous partial derivatives of the first n orders.

[6] The tensor analysis summation convention is used.

coördinate system can be chosen so that $\dot{x}^0 > 0$ along g. In our proofs we shall suppose that such a coördinate system has been chosen. Our results, however, are valid for all coördinate systems.

3. First necessary conditions. If we consider our problem as a non-parametric problem in tx-space we obtain at once the results of this section from the non-parametric case (XIII).

THEOREM 3.1. *If g is a minimizing arc then there exists a function $F = \lambda^0 f + \lambda^\beta(t)\phi_\beta$ such that the Euler-Lagrange equations*

$$(3.1) \qquad P_i = (d/dt)F_{\dot{x}^i} - F_{x^i} = 0, \qquad \phi_\beta = 0 \qquad (i = 0, 1, \cdots, n)$$

hold at each point of g. Moreover on g the equation[7]

$$(3.2) \qquad [F_{\dot{x}^i} dx^i]_1^2 + \lambda^0 d\theta = 0$$

is an identity in $d\alpha_h$ when the differentials dx^{i1}, dx^{i2}, $d\theta$ are expressed in terms of $d\alpha_h$. The multiplier λ^0 is a constant. The multipliers $\lambda^\beta(t)$ are continuous and the elements of the set λ^0, $\lambda^\beta(t)$ do not vanish simultaneously at any point of g.

It is readily proved that under the transformation of admissible coördinates (x) the functions P_i in equations (3.1) are covariant components of a vector and that the first member of the identity (3.2) is an invariant. It follows that the multipliers λ^0, $\lambda^\beta(t)$ belonging to g are invariants. Moreover the number of linearly independent multipliers of the form $\lambda^0 = 0$, $\lambda^\beta(t)$ with which a subarc g_1 of g satisfies the equations (3.1) is clearly a numerical invariant and is called the *order of anormality of g on g_1 relative to the equations* (3.1). Similarly the number of linearly independent multipliers of the form $\lambda^0 = 0$, $\lambda^\beta(t)$ with which g satisfies the conditions (3.1) and (3.2) is also a numerical invariant and is called the *order of anormality of g relative to the equations* (3.1) *and* (3.2). If there are no multipliers of the latter type g is said to be normal and the multiplier λ^0 in Theorem 3.1 can be taken to be unity and in this form the multipliers are unique.

We come now to the *necessary condition of Weierstrass* (see Graves, VIII).

THEOREM 3.2. *If g is a normal minimizing arc then at each element $(x, \dot{x}, \lambda)$ on g the inequality*

$$E(x, \dot{x}, \lambda, r) \geqq 0$$

[7] Here and elsewhere the symbol $[\ ]_1^2$ denotes the value of $[\ ]$ at the point 2 minus the value of $[\ ]$ at the point 1.

must hold for every set $(x, r) \neq (x, k\dot{x})$ $(k > 0)$ *which satisfies the equations* $\phi_\beta = 0$ *and gives the matrix* $\| \phi_{\beta r^i} \|$ *rank* m, *where*

$$E(x, \dot{x}, \lambda, r) = F(x, r, \lambda) - r^i F_{r^i}(x, \dot{x}, \lambda). \tag{3.3}$$

The *necessary condition of Clebsch* is described in the following theorem:

THEOREM 3.3. *If* g *is a normal minimizing arc then at each element* $(x, \dot{x}, \lambda)$ *on* g *the inequality*

$$F_{r^i r^j} \pi^i \pi^j \geqq 0 \qquad (i, j = 0, 1, \cdots, n) \tag{3.4}$$

must hold for all sets $(\pi) \neq (\tau\dot{x})$ *which satisfy the equations*

$$\phi_{\beta r^i} \pi^i = 0. \tag{3.5}$$

The Weierstrass E-function (3.3) is clearly an invariant. So also are the first members of the relations (3.4) and (3.5) provided that (π) is a contravariant vector.

We shall now define a new analogue of the Weierstrass function F_1. In order to do so we first note that the determinant

$$\begin{vmatrix} F_{r^i r^j} & \phi_{\beta r^i} \\ \phi_{\beta r^j} & 0 \end{vmatrix} \tag{3.6}$$

is identically zero along g, as follows readily from the equations

$$F_{r^i r^j} r^j = 0, \qquad \phi_\beta = r^j \phi_{\beta r^j} \tag{3.7}$$

which are found by differentiating the equations (2.3) for k and r^i and setting $k = 1$. Let $b_p{}^i(x, r)$ $(p = 1, \cdots, n)$ be a set of n contravariant vectors homogeneous of order zero in (r) and having its determinant $| b_p{}^i r^i |$ different from zero. By the *Weierstrassian function* F_1 is meant the determinant

$$F_1(x, r, \lambda) = \begin{vmatrix} b_p{}^i F_{r^i r^j} b_q{}^j & \phi_{\beta r^i} b_p{}^i \\ \phi_{\beta r^j} b_q{}^j & 0 \end{vmatrix}. \tag{3.8}$$

One readily verifies that the function F_1 is different from zero along g if and only if the matrix (3.6) has rank $n + m$ on g. It can be shown further that the function F_1 here defined differs from the function F_1 used hitherto by a non-vanishing factor (cf. VII, p. 104; X, pp. 112-113). Moreover the function F_1 here defined is an absolute invariant whereas the function F_1 used hitherto is a relative invariant of the second order.

By an *extremal arc* is meant a differentially admissible arc and a set of multipliers, not vanishing simultaneously

$$x^i = x^i(t), \quad \lambda^0, \quad \lambda^\beta = \lambda^\beta(t) \qquad (t^1 \leqq t \leqq t^2)$$

possessing continuous derivatives $\dot{x}^i$, $\ddot{x}^i$, $\dot{\lambda}^\beta$ and satisfying the equations (3. 1). An extremal is said to be *non-singular* if the function F_1 is different from zero at each point on it.

If the arc g under consideration satisfies the equations (3. 1) with a set of multipliers λ^0, $\lambda^\beta(t)$ and is such that $F_1(x, \dot{x}, \lambda)$ is different from zero on it, then g is an extremal arc. Moreover the parameter t can be chosen so that the functions $x^i(t)$ defining g are of class C^3 and the functions $\lambda^\beta(t)$ of class C^2 (cf. VII, p. 103).

In the sequel we shall assume that g is an extremal arc having $\lambda^0 = 1$ and satisfying the transversality condition (3. 2), *unless otherwise expressly stated.* It will be understood further that the parameter t has been chosen so as to give the maximum number of derivatives.

4. The second variation and the accessory minimum problem. By considering our problem as a non-parametric problem in tx-space with t^1 and t^2 fixed it is found that along g the second variation of the functional J takes the form

$$J_2(\eta, w) = b_{hl} w^h w^l + \int_{t^1}^{t^2} 2\omega(\eta, \dot{\eta})\, dt, \tag{4. 1}$$

where

$$2\omega = F_{x^i x^j} \eta^i \eta^j + 2F_{x^i r^j} \eta^i \dot{\eta}^j + F_{r^i r^j} \eta^i \dot{\eta}^j,$$
$$b_{hl} = \theta_{hl} + [F_{r^i} x^{is}_{hl}]_1^2,$$

the subscripts h, l in the second member of the last equation denoting differentiation with respect to α_h, α_l and setting $(\alpha) = (0)$. The variations $\eta^i(t)$ are assumed to be continuous and to have continuous derivatives except possible at a finite number of values of t on $t^1 t^2$ and to satisfy with the constants w^h the equations

$$\Phi_\beta(\eta) = \phi_{\beta x^i} \eta^i + \phi_{\beta r^i} \eta^i = 0, \tag{4. 2}$$

$$\eta^{is} - x_h^{is} w^h = 0 \qquad (s = 1, 2), \tag{4. 3}$$

where the coefficients of η^i, $\dot{\eta}^i$ in (4. 2) are evaluated along g and the subscript h denotes differentiation with respect to α_h at $(\alpha) = (0)$. Such variations (η, w) will be termed *admissible variations.* As was seen by Morse (X, p. 123) the variations η^i, x_h^{i1}, x_h^{i2} define respectively contravariant vectors. It follows that the condition (4. 3) is independent of the coördinate system

(x) used. Similarly the functions $\Phi_\beta(\eta)$ are invariants and the functional $J_2(\eta, w)$ is an invariant subject to the condition (4. 3), as can be seen from the way in which they were derived. The functions 2ω and b_{hl} however are not invariants, as one readily verifies.

The following necessary condition follows at once from the non-parametric case (XIII, p. 797):

THEOREM 4. 1. *If g is a normal minimizing extremal arc the second variation* (4. 1) *of J must satisfy along g the condition $J_2(\eta, w) \geqq 0$ for every set of admissible variations (η, w) having continuous second derivatives except possibly at a finite number of points on t^1t^2.*

The theorem just described suggests the study of the problem of minimizing the functional $J_2(\eta, w)$ in the class of admissible variations η^i, w^h. The problem is a non-parametric problem of Bolza and will be termed *the accessory minimum problem.* The analogues of equations (3. 1) and (3. 2) are the following:

$$L_i(\eta, \mu) \equiv (d/dt)\Omega_{\eta^i} - \Omega_{\eta^i} = 0, \qquad \Phi_\beta = 0, \tag{4. 4}$$

$$\zeta_i{}^2 x_h{}^{i2} - \zeta_i{}^1 x_h{}^{i1} + b_{hl} w^l = 0 \qquad (h, l = 1, \cdots, r), \tag{4. 5}$$

where

$$\Omega = \omega + \mu^\beta \Phi_\beta, \qquad \zeta_i = \Omega_{\dot\eta^i}.$$

Here we have taken the constant μ^0 analogous to λ^0 in (3. 1) to be unity, which is permissible since the accessory minimum problem can always be made normal. The equations (4. 4) are known as the *accessory differential equations,* the equations (4. 5) as the *secondary transversality conditions.* The extremals with $\mu^0 = 1$ for this problem will be called *secondary extremals.*

By an argument like that given by Morse (X, p. 123) it is readily seen that on g the functions $L_i(\eta, \mu)$ are covariant components of a vector and that the first members of equations (4. 5) are invariants subject to the conditions (4. 3). The multipliers μ^β belonging to a secondary extremal η^i, μ^β are therefore invariants, the functions η^i being contravariant as was seen above.

If $\tau(t)$ is a continuous function possessing a continuous derivative except possibly at a finite number of points on t^1t^2, then as is well known (VII, p. 114) the functions

$$\eta^i = \tau\dot x^i, \qquad \mu^\beta = \tau\lambda^\beta \tag{4. 6}$$

satisfy equations (4. 4). It follows that if η^i, μ^β is a secondary extremal so also is $\eta^i - \tau\dot x^i$, $\mu^\beta - \tau\dot\lambda^\beta$ provided the function $\tau(t)$ is of class C^2. It

should be noted also that the equations (4.4) are not all independent since the equation

$$\dot{x}^i L_i(\eta, \mu) = \lambda^\beta \Phi_\beta(\eta) \tag{4.7}$$

is an identity in the variables η^i, $\dot{\eta}^i$, $\ddot{\eta}^i$, μ^β, $\dot{\mu}^\beta$, as one readily verifies with the help of known consequences of the homogeniety conditions (2.3) such as equations (3.7).

The following formula will be useful in the next section:

$$J_2(\eta - \tau\dot{x}, w) = J_2(\eta, w) + [\tau^2 \dot{x}^i F_{x^i} + 2\tau F_{x^i}\eta_i]_1^2. \tag{4.8}$$

This relation holds for all functions η^i, w^h, τ having the continuity properties described in the above paragraphs and satisfying equations (4.2). In order to establish this relation we note that

$$\begin{aligned} 2\omega(\eta - \tau\dot{x}) &= 2\Omega(\eta - \tau\dot{x}) \\ &= 2\omega(\eta) + 2[\eta^i \Omega_{\eta^i}(\tau\dot{x}) + \dot{\eta}^i \Omega_{\dot{\eta}^i}(\tau\dot{x})] + 2\Omega(\tau\dot{x}) \end{aligned} \tag{4.9}$$

the multipliers μ^β in Ω being $\tau\lambda^\beta$. Noting that by virtue of equations (3.7) and (3.1)

$$\Omega_{\dot{\eta}^i}(\tau\dot{x}) = \tau(d/dt) F_{\dot{x}^i} = \tau F_{x^i} \tag{4.10}$$

and recalling that the functions (4.6) satisfy equations (4.4) it is found that

$$\eta^i \Omega_{\eta^i}(\tau\dot{x}) + \dot{\eta}^i \Omega_{\dot{\eta}^i}(\tau\dot{x}) = (d/dt)[\eta^i \Omega_{\dot{\eta}^i}(\tau\dot{x})] = (d/dt)(\tau F_{x^i}\eta^i). \tag{4.11}$$

Similarly from the well known identity

$$2\Omega = \eta^i \Omega_{\eta^i} + \dot{\eta}^i \Omega_{\dot{\eta}^i} + \mu^\beta \Omega_{\mu^\beta}$$

we find that

$$2\Omega(\tau\dot{x}) = (d/dt)(\tau^2 F_{x^i}\dot{x}^i). \tag{4.12}$$

The relation (4.8) now follows at once from the equation (4.9), (4.11), and (4.12).

5. The second variation in normal form. Since the determinant (3.6) is always zero it is clear that the accessory minimum problem described in the last section is singular. In the present section we shall set up an equivalent accessory minimum problem which is always non-singular whenever the function F_1 is different from zero along g. The method used is essentially a generalization of a method due to Weierstrass (See I, pp. 224-226) and seems to be particularly adaptable to the present paper (cf. II, IV, VII, XII).

Let $a_i^p(x)$ $(p = 1, \cdots, n)$ be a set of n linearly independent covariant vectors of class C^2 near g and satisfying the equations $a_i^p \dot{x}^i = 0$ along g. For

example we may choose these vectors so as to form on g a set of n mutually orthogonal vectors normal to g, as does Weierstrass in the two dimensional case. It is clear that the functions $u^p(t)$ defined along g by the equations

$$u^p = a_i{}^p \eta^i \qquad (p = 1, \cdots, n) \tag{5.1}$$

are invariants relative to transformations of admissible coördinates (x). Moreover the variations η^i and $\eta^i - \tau \dot{x}^i$ determine the same functions u^p. If now we select a covariant vector $a_i(x)$ such that along g we have $a_i \dot{x}^i \neq 0$, then there exist a set of n linearly independent contravariant vectors $b_p{}^i(x)$ of class C^2 with $a_i b_p{}^i = 0$ along g and such that the equations

$$\eta^i = b_p{}^i u^p + \tau \dot{x}^i, \qquad \tau = a_i \eta^i / a_j \dot{x}^j \tag{5.2}$$

hold along g, whenever the equations (5.1) are satisfied. It follows that a set of admissible variations $\eta^i(t)$, w^h determine a unique set of functions $u^p(t)$, w^h, $\tau(t)$ of class D^1 satisfying the equations

$$\Phi_\beta{}^0(u) = \Phi_\beta(b_p u^p + \tau \dot{x}) = \Phi_\beta(b_p u^p) = 0, \tag{5.3}$$

$$u^{ps} - c_h{}^{ps} w^h = 0, \qquad \tau^s = d_h{}^s w^h \tag{5.4}$$

where (s not summed)

$$\begin{aligned} &c_h{}^{ps} = a_i{}^{ps} x_h{}^{is}, \qquad d_h{}^s = a_i{}^s x_h{}^{is} / a_j{}^s \dot{x}^{js} \\ &x_h{}^{is} = b_p{}^{is} c_h{}^{ps} + d_h{}^s \dot{x}^{is} \qquad (s = 1, 2). \end{aligned} \tag{5.5}$$

Conversely every set of functions $u^p(t)$, w^h, $\tau(t)$ of class D^1 satisfying equations (5.3) and (5.4) determines by means of equations (5.2) a set of admissible variations $\eta^i(t)$, w^h. A set of functions $u^p(t)$, w^h of the type just described will be said to be a set of *normal admissible variations.*

Under the transformation (5.2) the second variation takes the form

$$J_2{}^0(u, w) = b_{hl}{}^0 w^h w^l + \int_{t^1}^{t^2} 2\omega^0(u, u)\, dt, \tag{5.6}$$

where

$$\begin{aligned} &2\omega^0 = 2\omega(b_p u^p), \\ &b_{hl}{}^0 = b_{hl} + [\dot{x}^i F_{x^i} d_h{}^s d_l{}^s + F_{x^i}(x_h{}^{is} d_l{}^s + x_l{}^{is} d_h{}^s)]_1{}^2, \end{aligned} \tag{5.7}$$

as is easily seen with the help of the formula (4.8). The expression (5.6) will be termed the *normal form of the second variation.* It is invariant subject to the conditions (5.4), although the functions $b_{hl}{}^0$, $2\omega^0$ are not invariants.

The following result is immediate:

LEMMA 5.1. *The functional $J_2(\eta, w)$ is positive (non-negative) for all admissible variations $(\eta, w) \neq (\tau\dot{x}, 0)$ if and only if the functional $J_2^0(u, w)$ is positive (non-negative) for all normal admissible variations $(u, w) \neq (0, 0)$.*

In view of this last lemma it is clear that the accessory minimum problem is equivalent to that of minimizing the functional $J_2^0(u, w)$ in the class of normal variations (u, w). This new problem will be termed the *normal accessory minimum problem.* It is of the type already studied by the author (XIII, p. 798) and its properties are known. It remains to interpret these properties in terms of the original problem. The analogues of equations (4.4) and (4.5) are the following

$$L_p{}^0(u, \rho) = (d/dt)\Omega_{\dot{u}^p}{}^0 - \Omega_{u^p}{}^0 = 0, \qquad \Phi_\beta{}^0 = 0 \tag{5.8}$$

$$v_p{}^2 c_k{}^{p2} - v_p{}^1 c_k{}^{p1} + b_{kl}{}^0 w^l = 0, \tag{5.9}$$

where

$$\Omega^0 = \omega^0 + \rho^\beta \Phi_\beta{}^0, \qquad v_p = \Omega_{\dot{u}^p}{}^0.$$

An extremal (u, ρ) for the problem just described will be termed a *normal secondary extremal.*

LEMMA 5.2. *Every secondary extremal (η, μ) determines a unique normal secondary extremal (u, ρ) and a function $\tau(t)$ of class C^2 such that the equations* (5.1) *and*

$$\eta^i = b_p{}^i u^p + \tau\dot{x}^i, \qquad \mu^\beta = \rho^\beta + \tau\lambda^\beta, \qquad \tau = a_i\eta^i/a_j\dot{x}^j \tag{5.10}$$

hold along g, and conversely. The corresponding values of ζ_i, v_p are connected by the formulas

$$\zeta_i = a_i{}^p v_p + \tau F_{x^i}, \qquad v_p = b_p{}^i(\zeta_i - \tau F_{x^i}). \tag{5.11}$$

This result will follow at once if we note that under the transformation (5.10) the functions L_i, $L_p{}^0$ of equations (4.4) and (5.8) are connected by the formulas

$$L_p{}^0 = b_p{}^i L_i, \qquad L_i = a_i{}^p L_p{}^0. \tag{5.12}$$

This result is readily established by substitution.

In view of equations (5.12) it is clear that the functions $L_p{}^0(u, \rho)$ are invariants. Similarly the first members of equations (5.9) are invariants subject to the conditions (5.4) since under the transformation (5.10) they are equal to the first members of equations (4.5), as one readily verifies with the help of equations (5.5), (5.7), and (5.11).

Suppose now that g is non-singular. The determinant (3.8) is then different from zero along g and the normal secondary extremals (u, ρ) are therefore also non-singular. The following result is an immediate consequence of Lemma 5.2 and known theorems concerning the solutions of equations (5.8).

THEOREM 5.1. *If the extremal g is non-singular then the rank of the matrix*

$$\left\| \begin{matrix} \eta_j{}^i & \dot{x}^i \\ \zeta_{ij} & F_{x'^i} \end{matrix} \right\|, \qquad \zeta_{ij} = \Omega_{\dot{\eta}^i}(\eta_j, \dot{\eta}_j, \mu_j) \tag{5.13}$$

formed for p secondary extremals $\eta_j{}^i$, $\mu_j{}^\beta$ $(j = 1, \cdots, p)$ is the same at each point on t^1t^2. Moreover every secondary extremal (η, μ) is expressible linearly in the form

$$\eta^i = \eta_j{}^i a^j + \tau \dot{x}^i, \qquad \mu^\beta = \mu_j{}^\beta a^j + \tau \lambda^\beta \tag{5.14}$$

in terms of $p = 2n$ secondary extremals $\eta_j{}^i$, $\mu_j{}^\beta$ whose matrix (5.13) has rank $2n + 1$, the a's being constants and the function $\tau(t)$ being of class C^2. The corresponding value of ζ_i is

$$\zeta_i = \zeta_{ij} a^j + \tau F_{x'^i}. \tag{5.15}$$

With the help of the last theorem and Lemma 5.2 one can readily establish analogues of necessary conditions for the second variation to be non-negative as given by the author for the non-parametric case (XIII, pp. 801-803).

We come now to the notion of conjugate points. A value $t^3 \neq t^1$ will be said to define a point 3 *conjugate* to 1 on g if there exists a secondary extremal (η, μ) having $\eta^i(t^1) = \eta^i(t^3) = 0$ and $(\eta) \not\equiv (\tau \dot{x})$ on t^1t^3. By virtue of the equations (5.10) this is equivalent to the condition that there exists a normal secondary extremal (u, ρ) with $u^p(t^1) = u^p(t^3) = 0$ and $(u) \not\equiv (0)$ on t^1t^3. The following theorem is an immediate consequence of equations (5.10) and a theorem of Hestenes (XIII, p. 804).

THEOREM 5.2. *Suppose the extremal g is non-singular. If $\eta_j{}^i$, $\mu_j{}^\beta$ $(j = 1, \cdots, 2n)$ form $2n$ secondary extremals whose matrix (5.13) has rank $2n + 1$ then a value $t^3 \neq t^1$ defines a point 3 conjugate to the point 1 on g if and only if the matrix*

$$\left\| \begin{matrix} \eta_j{}^i(t^3) & \dot{x}^i(t^3) & 0 \\ \eta_j{}^i(t^1) & 0 & \dot{x}^i(t^1) \end{matrix} \right\|$$

has rank $r < 2n - q + 2$, where q is the order of anormality of g in its subarc determined by t^1t^3.

6. The accessory boundary value problem. Let $K_{ij}(x, r)$ be a set of

functions of class C^1 in a neighborhood of the values $(x, r) = (x, \dot{x})$ on g, and such that along g we have $K_{ij} = K_{ji}$ and

$$K_{ij}\eta^i\eta^j > 0 \qquad (\eta) \neq (\tau\dot{x}), \qquad K_{ij}r^j = 0. \tag{6.1}$$

Moreover these functions are assumed to be positively homogeneous of the first order in (r) and to be the components of a covariant tensor of the second order. For example, we can in general select $K_{ij} = F_{r^ir^j}$. By the accessory boundary value problem is meant the system of equations

$$\begin{aligned} &M_i(\eta, \mu, \sigma) = L_i(\eta, \mu) + \sigma K_{ij}\eta^j = 0, \qquad \Phi_\beta = 0, \\ &\eta^{is} - x_h{}^{is}w^h = 0 \qquad\qquad (i, j = 0, 1, \cdots, n;\ s = 1, 2) \\ &\zeta_i{}^2x_h{}^{i2} - \zeta_i{}^1x_h{}^{i1} + b_{hl}w^l = 0 \qquad\qquad (h, l = 1, \cdots, r), \end{aligned} \tag{6.2}$$

where $L_i(\eta, \mu)$, Φ_β, ζ_i are the functions appearing in equations (4.4) and (4.5). By a *characteristic root* is meant a value of σ for which there exists a solution (η, μ, w) of these equations having continuous derivatives $\dot{\eta}^i$, $\ddot{\eta}^i$, $\dot{\mu}^\beta$ and with $(\eta) \neq (\tau\dot{x})$. In view of our remarks in Section 4 and the choice of the functions K_{ij} it is clear that the system (6.2) is independent of the coördinate system (x) used. It follows that the characteristic roots σ are invariants.

The following theorem can be established by the usual method (See, for example, VI, p. 524).

THEOREM 6.1. *If the second variation $J_2(\eta, w)$ of the functional J is non-negative along g for every set of admissible variations (η, w), then there can be no negative characteristic roots of the accessory boundary value problem* (6.2).

The end conditions (1.3) will be said to be *regular* on g if the matrix obtained by deleting the last two columns of the matrix

$$\left\| \begin{matrix} x_h{}^{i1} & \dot{x}^i(t^1) & 0 \\ x_h{}^{i2} & 0 & \dot{x}^i(t^2) \end{matrix} \right\| \tag{6.3}$$

has rank r. The arc g will be said to satisfy the *non-tangency condition* in case the last two columns are not linearly dependent on the first r columns. The matrix (6.3) accordingly has rank $r + 2$ if and only if the end conditions are regular and the non-tangency condition holds on g.

The extremal g will be said to satisfy the *Clebsch S-condition* if the relation (3.4) with the equality sign excluded holds along g for every set $(\pi) \neq (\tau\dot{x})$ satisfying the equations (3.5).

We can now prove the theorem:

THEOREM 6.2. *Suppose the end conditions are regular and the non-tangency condition holds on the extremal g. If g is non-singular, then the second variation $J_2(\eta, w)$ of J is positive (non-negative) for every set of admissible variations $(\eta, w) \neq (\tau\dot{x}, 0)$ if and only if the Clebsch S-condition holds and all the characteristic roots of the accessory boundary value problem are positive (non-negative).*

To prove this we first note that under the transformation (5.10) the system of equation (6.2) are equivalent to the system

$$
\begin{aligned}
& M_p{}^0(u, \rho, \sigma) = L_p{}^0(u, \rho) + \sigma K_{pq}{}^0 u^q = 0, \qquad \Phi_\beta{}^0 = 0, \\
& u^{ps} - c_h{}^{ps} w^h = 0 \qquad\qquad (p, q = 1, \cdots, n;\ s = 1, 2), \\
& v_i{}^2 c_h{}^{i2} - v_i{}^1 c_h{}^{i1} + b_{hl}{}^0 w^l = 0 \qquad (h, l = 1, \cdots, r),
\end{aligned} \tag{6.4}
$$

where the functions $L_p{}^0(u, \rho)$, $\Phi_\beta{}^0$, v_p, $c_h{}^{i1}$, $c_h{}^{i2}$ are those appearing in equations (5.8), (5.9) and

$$
K_{pq}{}^0 = b_p{}^i K_{ij} b_q{}^j, \qquad K_{ij} = a_i{}^p K_{pq}{}^0 a_j{}^q.
$$

We may suppose further that the functions $b_p{}^i(x)$ have been chosen so that $K_{pq}{}^0$ is equal to the Kronecker delta along g. The system of equations (6.4) is then the boundary value problem usually associated with the functional $J_2{}^0(u, w)$. Its characteristic roots are defined to be the values of σ for which there exists a solution (u, ρ, w) having continuous derivatives $\dot{u}^p$, $\ddot{u}^p$, $\dot{\rho}^\beta$ and with $(u) \neq (0)$. It is clear that the functions $M_i, M_p{}^0$ satisfy equations (5.12) with L replaced by M. It follows that a value σ is a characteristic root of the boundary value problem (6.2) if and only if it is a characteristic root of the system (6.4). Moreover the analogue of Theorem 6.2 is true for the functional $J_2{}^0(u, w)$ in the class of normal admissible variations (u, w), as has been shown by the author (XIII, pp. 812-813). The theorem now follows from Lemma 5.1.

7. Sufficient conditions for relative minima. In this section we shall show that sufficient conditions as well as necessary conditions can be derived with the help of known theorems for the non-parametric problem. To this end we introduce the notion of a Mayer field.

A *Mayer field* is a region $\mathcal{F}$ in $\mathcal{R}$ together with a set of slope functions $r^i(x)$ and multipliers $\lambda^\beta(x)$ with the following properties:

(a) The slope functions $r^i(x)$ are the contravariant components of a vector of class C' and with $(r) \neq (0)$. The multipliers $\lambda^\beta(x)$ are invariant functions of class C'.

(b) The elements (x, r) satisfy the equations $\phi_\beta = 0$.

(c) The Hilbert integral

$$I^* = \int F_{r^i}(x, r, \lambda)\, dx^i$$

formed for these functions and $\lambda^0 = 1$ is independent of the path in $\mathfrak{F}$.

It is well known that a solution $x^i = x^i(t)$ of the equations $\dot{x}^i - r^i(x) = 0$ together with the multipliers $\lambda^\beta[x(t)]$ define an extremal, called an *extremal of the field* $\mathfrak{F}$.

If g is an extremal of a field $\mathfrak{F}$ then the conditions (3.2) for g can be written in the form

$$[dI^*]_1^2 + d\theta \equiv 0. \tag{7.1}$$

In view of this fact it follows readily that the second differential

$$[d^2I^*]_1^2 + d^2\theta \tag{7.2}$$

is a quadratic form in $d\alpha_h$ whenever the relation (7.1) holds identically in $d\alpha_h$. The Weierstrass E-function $E(x, r, \lambda, \dot{x})$ is defined by equations (3.3).

By an argument like that used by Hestenes in the non-parametric case (XIII, pp. 805-806) we can prove the following theorem:

THEOREM 7.1. *If g is an extremal of a Mayer field at each point of which the inequality*

$$E[x, r(x), \lambda(x), \dot{x}] \geqslant 0 \tag{7.3}$$

holds for every set $(x, \dot{x}) \neq (x, kr)$ $(k > 0)$ satisfying equations $\phi_\beta = 0$ and if the endpoints of g and the set $(\alpha) = (0)$ are such that the equation (7.1) is an identity in $d\alpha_h$ and the quadratic form (7.2) is positive definite, then the arc g affords a proper minimum to the functional J relative to admissible arcs C in $\mathfrak{F}$ with sets (α) near $(\alpha) = (0)$.

An extremal arc will be said to satisfy the *Weierstrass S-condition* if at each element $(x, \dot{x}, \lambda)$ on it the inequality

$$E(x, \dot{x}, \lambda, r) > 0 \tag{7.4}$$

holds for every set $(x, r) \neq (x, kx)$ $(k > 0)$ satisfying the equations $\phi_\beta = 0$. The *Clebsch S-condition* has been defined in the paragraph preceding Theorem 6.2.

LEMMA 7.1. *If an extremal g satisfies the Weierstrass and Clebsch S-conditions then the inequality (7.4) holds for all sets $(x, \dot{x}, \lambda)$ in a neigh-*

borhood of those belonging to g and for every set $(x, r) \neq (x, k\dot{x})$ $(k > 0)$ satisfying the equations $\phi_\beta = 0$.

The proof is well known and can be made by the arguments like those used by Morse (X, p. 121) for the case in which there are no differential side conditions.

We come now to the following important result:

THEOREM 7.2. *In order that an admissible arc g without corners and not intersecting itself afford a proper strong relative minimum to the functional J it is sufficient that there exists a set of multipliers $\lambda^0 = 1$, $\lambda^\beta(x)$ with which g satisfies the Euler-Lagrange equations (3.1), the transversality condition (3.2), the Weierstrass S-condition, the Clebsch S-condition, and that the second variation (4.1) of J be positive along g for every set of admissible variations $(\eta, w) \neq (\tau\dot{x}, 0)$.*

In order to prove this theorem we shall construct an auxiliary non-parametric problem of Bolza as follows: Let a neighborhood of g be represented by a single admissible coördinate system (x) with $\dot{x}^0(t) > 0$ along as described in Lemma 2.1. The coördinates $(x^0, x^1, \cdots, x^n)$ will also be termed $(x, y_1, \cdots, y_n)$. Since the Clebsch S-condition implies that g is a non-singular extremal, the function $\dot{x}^0(t)$ is of class C^3 for admissible parameters t. We can accordingly choose $t = x^0 = x$ as the parameter t defining g. The equations for g can then be written in the form

$$y_p = y_p(x) \qquad (x_1 \leqq x \leqq x_2;\ p = 1, \cdots, n). \tag{7.5}$$

Moreover the functions ϕ, f can be written in the form

$$\phi_\beta(x, y_1, \cdots, y_n, 1, y'_1, \cdots, y'_n) = \phi_\beta{}^0(x, y, y'),$$
$$f(x, y, \cdots, y,\ 1, y', \cdots, y'\ ) = f^0(x, y, y').$$

This suggests the problem of minimizing the functional

$$J^0 = \theta(\alpha) + \int_{x_1}^{x_2} f^0(x, y, y')\, dx$$

in the class of arcs (7.5) and sets (α) which satisfy the differential equations and end conditions

$$\phi_\beta{}^0(x, y, y') = 0 \qquad (\beta = 1, \cdots, m \leq n),$$
$$x_s = x_s(\alpha), \qquad y_p(x_s) = y_{ps}(\alpha) \qquad (s = 1, 2),$$

where

$$x_s(\alpha) = x^{0s}(\alpha), \qquad y_{ps}(\alpha) = x^{ps}(\alpha) \qquad (p = 1, \cdots, n).$$

This problem is of the type already studied by the author. Clearly g is also

an extremal for this new problem, its multipliers being $\lambda^0 = 1$, $\lambda^\beta(x)$. The analogue of the Clebsch S-conditions also holds along g. Moreover if in Section 5 we select the functions $a_i{}^p$ so that for the coördinate system here used the equations (5. 1) take the form

$$u^p = \eta^p - \dot{x}^p(x^0)\eta^0 \qquad (p = 1, \cdots, n)$$

and the functions a_i so that $(a_0, a_1, \cdots, a_n) = (1, 0, \cdots, 0)$ then by setting $t = x$ it is found that the functional (5. 6) is the second variation of the functional J^0, that the equations (5. 3) are the equations of variation of the equations $\phi_\beta{}^0 = 0$, and that the equations (5. 4) are the secondary end conditions for the problem just described. It follows from Lemma 5. 1 that the second variation $J_2{}^0(u, w)$ is positive for every non-null set of admissible variations (u, w).

As a second step in the proof of Theorem 7. 2 we note that it is sufficient to prove the theorem for the case in which the end conditions (1. 3) and the function $\theta(\alpha)$ are of the form

$$x^{i1} = x^{i1}(\alpha^1, \cdots, \alpha^\rho), \qquad x^{i2} = x^{i2}(\alpha^{\rho+1}, \cdots, \alpha^r),$$
$$\theta(\alpha) = \theta^1(\alpha^1, \cdots, \alpha^\rho) - \theta^2(\alpha^{\rho+1}, \cdots, \alpha^r).$$

The proof of this statement is like that given by the author for the non-parametric case (XIII, pp. 815-816). The end conditions for the auxiliary non-parametric described in the preceding paragraph will also be in this separated form. It follows from the Clebsch S-condition and the positiveness of the second variation that there exists for this non-parametric problem a Mayer field $\mathfrak{F}$ such that the analogue of equation (7. 1) holds and that the quadratic form analogous to (7. 2) is positive definite. But when these conditions are interpreted in terms of the parametric problem it is found that $\mathfrak{F}$ is also a Mayer field for the parametric problem and that the condition (7. 1) holds and that the quadratic form (7. 2) is positive definite. By taking $\mathfrak{F}$ sufficiently small the condition (7. 3) will also be satisfied by virtue of the Weierstrass S-condition and Lemma 7. 1. Theorem 7. 2 now follows Theorem 7. 1.

Combining Theorems 6. 2 and 7. 2 we obtain the following result:

THEOREM 7. 3. *Suppose the end conditions are regular and the non-tangency condition holds on an admissible arc g without corners and not intersecting itself. If there exists a set of multipliers $\lambda^0 = 1$, $\lambda^\beta(t)$ with which g satisfies the Euler-Lagrange equations* (3. 1), *the transversality condition* (3. 2), *the Weierstrass S-condition, the Clebsch S-condition, and if all the*

characteristic roots of the accessory boundary value problem are positive, then g affords a proper strong relative minimum to J.

Similarly the analogues of the sufficient conditions of "Theorems 9. 1, 9. 3, 9. 4" of Hestenes (XIII, cf. XVI) and also those of Morse (XIV) and Reid (XV) can be established without difficulty.

HARVARD UNIVERSITY.

REFERENCES.

I. Bolza, *Vorlesungen über Variationsrechnung*, 1909.

II. Bliss, "Jacobi's condition for problems of the calculus of variations in parametric form," *Transactions of the American Mathematical Society*, vol. 17 (1916), pp. 195-206.

III. Eshleman, *The Lagrange Problem in Parametric Form*, Dissertation, The University of Chicago, 1922.

IV. Graves, "Discontinuous solutions in space problems of the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 1-28.

V. Bliss, "The problem of Lagrange in the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 673-744.

VI. Morse, "Sufficient conditions in the problem of Lagrange with variable end points," *American Journal of Mathematics*, vol. 53 (1931), pp. 517-596.

VII. Hefner, "The condition of Mayer for discontinuous solutions of the Lagrange problem," *Contributions to the Calculus of Variations* (1931-2), The University of Chicago Press, pp. 95-130.

VIII. Graves, "On the Weierstrass condition for the problem of Bolza in calculus of variations," *Annals of Mathematics* (2), vol. 33 (1932), pp. 747-752.

IX. Tucker, "On tensor invariance in the calculus of variations," *Annals of Mathematics* (2), vol. 35 (1934), pp. 341-350.

X. Morse, "The calculus of variations in the large," *American Mathematical Society Colloquium Publications*, vol. 18, New York (1934).

XI. Reid, "Analogues of the Jacobi condition for the problem of Mayer in the calculus of variations," *Annals of Mathematics*, vol. 35 (1934), pp.

XII. Hestenes, "A note on the Jacobi condition for parametric problems in the calculus of variations," *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 297-302.

XIII. Hestenes, "Sufficient conditions for the problem of Bolza in the calculus of variations," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 793-818.

XIV. Morse, "Sufficient conditions in the problem of Lagrange without assumptions of normalcy," *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 147-160.

XV. Reid, "The theory of the second variation for the non-parametric problem of Bolza," *American Journal of Mathematics*, vol. 57 (1935), pp. 573-586.

XVI. Hestenes, "On sufficient conditions in the problems of Lagrange and Bolza." To appear soon in the *Annals of Mathematics*.

ON EXPANSIONS IN TERMS OF A CERTAIN GENERAL CLASS OF FUNCTIONS.

By W. T. MARTIN.

Introduction. The object of this paper is to define certain generalizations of Appell polynomials and to study the problem of expanding a given function in a series of these generalized functions. A set of *Appell polynomials* $\{P_n(x)\}$ is defined by $e^{xz}\sum_0^\infty p_n z^n \sim \sum_0^\infty P_n(x)z^n$ where the series $\Sigma p_n z^n$ may or may not have a region of convergence. The functions $\{P_n(x)\}$ are such that $P_n(x)$ is a polynomial of degree n in x and $P'_n(x) = P_{n-1}(x)$, $n > 0$. Polynomials of this character were introduced by Appell and they, together with certain generalizations of them, have been studied by various writers.[1] When $\Sigma p_n z^n = z/(e^z - 1)$, the polynomials $\{n!\, P_n(x)\}$ are the *Bernoulii polynomials* $\{B_n(x)\}$. The *Bernoulli-Hurwitz functions* $B_{n,r}(x)$ $(n, r = 0, 1, 2, \cdots)$, are defined by

$$ze^{xz}/(e^z - 1) = \sum_{n=-\infty}^{\infty} B_{n,r}(x)z^n/n!, \qquad 2\pi r < |z| < 2\pi(r+1).$$

Carmichael has developed the theory of the expansion of functions in series of Bernoulli polynomials and of Bernoulli-Hurwitz functions.[2] Many of the ideas of the present paper were suggested by a study of his paper.

Replacing the function e^{xz} by a function $E(x, z)$ of a given character and denoting by $g(z)$ an analytic function having specified zeros, we define functions $A_{n,r}(x)$ by means of the Laurent's developments of $E(x, z)/g(z)$ in annular rings. In part I we study convergence properties of series of the sort $\sum_0^\infty c_n A_{n,r}(x)$, where r is a fixed non-negative integer.

In part II we study the properties of the functions represented by convergent series of the above sort and determine necessary and sufficient conditions on a given function in order that it be expressible in a series of this sort.

[1] P. Appell, *Ann. Sci. Norm. Sup.*, ser. 2, vol. 9 (1880), pp. 119-144; H. Léauté, *Journal de Mathématiques*, ser. 3, vol. 7 (1881), pp. 185-200; G. H. Halphen, *Bulletin des Sciences Mathématiques*, ser. 2, vol. 5 (1881), pp. 462-468; A. Angelesco, *Comptes Rendus*, vol. 176 (1923), pp. 275-278; I. M. Sheffer, *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 261-280; I. M. Sheffer, *American Journal of Mathematics*, vol. 53 (1931), pp. 15-38.

[2] R. D. Carmichael, *Annals of Mathematics*, ser. 2, vol. 34 (1933), pp. 349-378.

In part III we introduce an operator B which is such that $BA_{n,r}(x) = A_{n-1,r}(x)$. The problem of expanding a given function in a series of the above sort may also be viewed as a problem of solving a certain functional equation. We give existence theorems for these associated functional equations. When $E(x,z) = e^{xz}$ the function $BF(x)$ is the derivative $F'(x)$ of functions $F(x)$ of finite exponential type. For this case the theory furnishes solutions of exponential type of differential equations of infinite order, yielding results previously obtained by Pincherle and Perron.[3] As a corollary we obtain the solutions of a class of infinite systems of linear equations in an infinite number of unknowns treated previously (by different methods) by Perron.[4] When $E(x,z) = (1+z)^x$ the function $BF(x)$ is the difference $\Delta F(x) \equiv F(x+1) - F(x)$ of a certain class of integral functions.

I. Convergence Theory.

1.1. Definition of the functions $A_{n,r}(x)$. Let $u_0(x), u_1(x), \cdots$, be any infinite sequence of integral functions possessing the following two properties:

1°. $\limsup_{n=\infty} |u_n(x)|^{1/n} \leqq 1/f$ $(0 < f \leqq \infty)$, for all x;

2°. if a function $F(x)$, representable by a series of the form $\sum_0^\infty e_n u_n(x)$ for which

$$\limsup_{n=\infty} |e_n|^{1/n} < f,$$

is identically zero, then $e_n = 0$, $n = 0, 1, \cdots$.

Examples of such sequences with interesting connections are:

$$\{x^n/n!\}, \qquad \{x^n/(\lambda_0\lambda_1\cdots\lambda_n)\}, \qquad \{x(x-1)\cdots(x-n+1)/n!\},$$
$$\{x(x-\lambda_1)\cdots(x-\lambda_{n-1})/(\lambda_1\cdots\lambda_n)\},$$

where $\lambda_n \neq 0$, $\lim |\lambda_n| = \infty$ and in the fourth example $|\lambda_1| \leqq |\lambda_2| \leqq \cdots$.

Let us form the function $E(x,z) = \sum_0^\infty u_n(x) z^n$ which, we see, is analytic for all $|z| < f$ and for all x. Let $g(z) = \sum_0^\infty \beta_n z^n$, $\beta_0 \neq 0$, be an analytic function regular for $|z| < \rho$ $(0 < \rho \leqq f)$. Let $g(z)$ have zeros[5] of multiplicities s_m at points α_m, $|\alpha_m| < \rho$, $(m = 1, 2, \cdots)$.

---

[3] S. Pincherle, *Acta Mathematica*, vol. 48 (1926), pp. 279-304 (first published in 1888); O. Perron, *Mathematische Annalen*, vol. 84 (1921), pp. 31-42.

[4] O. Perron, *ibid.*, pp. 1-15. See in particular Satz 1.

[5] We will carry through the theory based upon the assumption that $g(z)$ has an

As a matter of notation we assume that the zeros α_m of $g(z)$ are arranged according to increasing absolute values and that those of the same absolute value are arranged according to increasing arguments, where the argument θ_m of α_m is taken to be in the range $0 \leqq \theta_m < 2\pi$. Let there be $k(1)$ zeros of smallest absolute value γ_1, $k(2)$ of next smallest absolute value γ_2, etc. Let us denote by $\mu(r)$ the sum $k(1)+\cdots+k(r)$. We define $\gamma_0 = k(0) = \mu(0) = 0$. Let C_r be a circle of radius C'_r, $\gamma_r < C'_r < \gamma_{r+1}$, about the origin as center.

Let us consider the expansions [6]

$$(1.1) \qquad \frac{E(x,z)}{g(z)} = \sum_{n=-\infty}^{\infty} A_{n,r}(x) z^n, \quad \gamma_r < |z| < \gamma_{r+1}, \quad (r = 0, 1, \cdots).$$

We see that the $A_{n,r}(x)$ are expressible as

$$(1.2) \qquad A_{n,r}(x) = \frac{1}{2\pi i} \int_{C_r} \frac{E(x,z)}{g(z)} \frac{dz}{z^{n+1}}.$$

Denoting by $R_n(x, \alpha_m)\alpha_m^{-n-1}$ the residue of $E(x,z)^{-n-1}/g(z)$ at $z = \alpha_m$, we see that $R_n(x, \alpha_m)$ has the form

$$(1.3) \qquad R_n(x, \alpha_m) = \sum_{k=0}^{s_m-1} \frac{(-1)^k}{\alpha_m^k} \binom{n+k}{k} p_{m,k}(D_m) E(x,z), \quad (n = 0, 1, \cdots),$$

where $p_{m,k}(t)$ is a polynomial of degree $s_m - 1 - k$ in t and D_m is an operator defined by

$$(1.4) \qquad D_m{}^j H(x,z) \equiv \frac{\partial^j}{\partial z^j} H(x,z) \Big|_{z=\alpha_m}.$$

1.2. Convergence theorems. Let us contemplate series of the form

$$(1.5) \qquad \sum_{n=0}^{\infty} c_n A_{n,r}(x),$$

where r is a fixed non-negative integer. Let us assume that this series converges [converges absolutely] for h points $x_1, \cdots, x_h$ for which the determinant Δ is non-vanishing. The determinant Δ contains h rows whose j-th one we indicate:

---

infinite number of zeros. No essential modification is necessary for the case where $g(z)$ has a finite number of zeros $a_1, \cdots, a_R$, $R \geqq 1$.

[6] The functions $P_{n,r}(x)$, $(n, r = 0, 1, \cdots)$, generated by $exz/g(z)$ are Appell polynomials $(r = 0)$ and associated Appell functions. Bird has introduced the functions $P_{-n,r}(x)$, $(n, r = 1, 2, \cdots)$, as the coefficients of the Laurent's expansion of $exz/g(z)$ in a certain annular ring. See M. T. Bird, *On generalizations of sum formulas of the Euler-Maclaurin type*, Illinois dissertation (1934).

$$\Delta \equiv \left| \delta^0_{\mu(r)+1,j} \cdots \delta^{\xi(1)}_{\mu(r)+1,j}\ \delta^0_{\mu(r)+2,j} \cdots \delta^{\xi(2)}_{\mu(r)+2,j} \cdots \delta^0_{\mu(r+1),j} \cdots \delta^{\xi(k(r+1))}_{\mu(r+1),j} \right|,$$

where

$$h = \sum_{m=\mu(r)+1}^{\mu(r+1)} s_m, \qquad \delta^k{}_{m,j} \equiv D_m{}^k E(x_j, z), \qquad \xi(q) + 1 = s_{\mu(r)+q}.$$

Let us assume temporarily that

$$\limsup_{n=\infty} |\, c_n \,|^{1/n} \leqq \gamma_{r+1}. \tag{1.6}$$

From (1.1) we see that

$$\limsup_{n=\infty} |\, A_{n,r}(x) \,|^{1/n} = 1/\gamma_{r+1}, \tag{1.7}$$

and hence in view of (1.6) each of the series

$$\sum_{n=0}^{\infty} c_n A_{n,r+p}(x), \qquad (p = 1, 2, \cdots), \tag{1.8}$$

converges absolutely and uniformly in every finite region. Forming (1.8) for $p = 1$ and subtracting from it the series (1.5), we see that the difference

$$\sum_{n=0}^{\infty} \frac{c_n}{\gamma_{r+1}^{n+1}} \sum_{m=\mu(r)+1}^{\mu(r+1)} e^{-i(n+1)\theta_m} \sum_{k=0}^{s_m-1} \frac{(-1)^k}{\alpha_m{}^k} \binom{n+k}{k} p_{m,k}(D_m) E(x, z) \tag{1.9}$$

converges [converges absolutely] for $x = x_1, \cdots, x_h$.

The determinant Δ_1 obtained from Δ by replacing $\delta^k{}_{m,j}$ by $p_{m,k}(D_m)E(x_j,z)$ is non-vanishing since it is obtainable from Δ by means of a finite number of elementary transformations of determinants. Since Δ_1 is different from zero the convergence [absolute convergence] of the series (1.9) for $x = x_1, \cdots, x_h$ implies the convergence [absolute convergence] of each of the series

$$\sum_{n=0}^{\infty} \frac{c_n}{\gamma_{r+1}^{n+1}} \binom{n+k}{k} e^{-i(n+1)\theta_m} \qquad (k = 0, \cdots, s_m - 1;\ m = \mu(r) + 1, \cdots, \mu(r+1)). \tag{1.10}$$

Hence the series (1.9) converges [converges absolutely] and uniformly in every finite region. Using the absolute and uniform convergence of the series (1.8) for $p = 1$, we obtain uniform [absolute and uniform] convergence of the series (1.5) in every finite region.

The preceding work depends upon the assumption that (1.6) holds. We show now that (1.6) must hold if the series (1.5) converges at the h points

$x_1, \cdots, x_h$. Assume for the moment that (1.6) does not hold.[7] Let η be a constant such that

$$\gamma_{r+1} < \eta < \gamma_{r+2}, \qquad \eta < \limsup_{n=\infty} |\, c_n \,|^{1/n}.$$

Consider the series (1.5) formed with c_n replaced by $c_n\zeta_n$ where

$$\zeta_0 = 1, \quad \zeta_n = 1/n^2 \text{ if } |\, c_n \,| \leqq \eta^n, \quad \zeta_n = \eta^n/(n^2 \,|\, c_n \,|) \text{ if } |\, c_n \,| > \eta^n.$$

It surely converges absolutely for $x = x_1, \cdots, x_h$ and furthermore we have

$$\limsup_{n=\infty} |\, c_n\zeta_n \,|^{1/n} = \eta.$$

From these facts we see that the series (1.8) formed with c_n replaced by $c_n\zeta_n$ converge absolutely and uniformly in every finite region. Reasoning as before, we have convergence for each of the series (1.10) with c_n replaced by $c_n\zeta_n$. But

$$\lim_{n=\infty} |\, c_n\zeta_n\gamma_{r+1}^{-n-1} \,| \neq 0.$$

So we are led to a contradiction and (1.6) must hold.

We have proved the following theorem and corollary:

THEOREM 1.1. *If the series* (1.5) *converges* [*converges absolutely*] *for* h *points* $x_1, \cdots, x_h$ *for which the determinant* Δ *is non-vanishing, then in every finite region of the* x*-plane each of the series* (1.5) *and* (1.8) *converges* [*converges absolutely*] *and uniformly.*

COROLLARY 1.1. *Under the hypotheses of the preceding theorem the series*

$$\phi(z) = \sum_{n=0}^{\infty} c_n z^{-n-1} \tag{1.11}$$

converges [*converges absolutely*] *for* $z = \alpha_m$ $(m = \mu(r) + 1, \cdots, \mu(r+1))$, *and hence the series* (1.11) *converges absolutely for* $|\, z \,| > \gamma_{r+1}$. *Indeed, each of the series* (1.10) *converges* [*converges absolutely*].

Let R be any finite region of the x-plane. We want to show that there exists in R a set of h points $x_1, \cdots, x_h$ for which Δ is non-vanishing. We use a special case of the following lemma:

LEMMA 1.1. *If* $a_1, \cdots, a_p$ *are any distinct non-zero constants,* $|\, a_k \,| < f$ $(k = 1, \cdots, p)$, *and if*

$$\sum_{k=1}^{p} \sum_{s=0}^{q} c_{sk} (\partial^s/\partial z^s) E(x, z) \,|_{z=a_k} \equiv 0,$$

then $c_{sk} = 0$, $(s = 0, \cdots, q;\ k = 1, \cdots, p)$.

[7] The fundamental ideas of this argument are due to H. J. Miles. I am also indebted to Miles for much of the notation of this part.

The truth of this lemma is an immediate consequence of the property 2° of the functions $\{u_n(x)\}$. Using this lemma and making an easy induction argument, we obtain the following lemma:

LEMMA 1.2. *In every region R of the x-plane there exists a set h points $x_1, \cdots, x_h$ for which the determinant Δ is non-vanishing.*

Theorem 1.2 follows from the preceding lemma.

THEOREM 1.2. *The hypotheses of Theorem* 1.1 *are satisfied whenever there is any region in which the series* (1.5) *converges* [*converges absolutely*].

II. EXPANSIONS IN TERMS OF THE FUNCTIONS $A_{n,r}(x)$.

2.1. A class of integral functions. Let $F(x)$ be any integral function enjoying the property that it has an expansion of the form

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x) \tag{2.1}$$

with

$$\limsup_{n=\infty} |\, e_n \,|^{1/n} = q, \qquad 0 \leqq q < f. \tag{2.2}$$

We shall, in such a case, say that $F(x)$ is of *sort* $\{u_n\}$, *type* q. The proof of the following lemma relating to functions of sort $\{u_n\}$, type q is immediate.

LEMMA 2.1. *A necessary and sufficient condition that a function $F(x)$ be of sort $\{u_n\}$, type not exceeding q, is that it be expressible in the form*

$$F(x) = (1/2\pi i) \int_C E(x, z) a(z)\, dz,$$

where $a(z)$ is analytic for z in the annular ring $q < |\, z \,| \leqq s \leqq \infty$ and C is a circle of radius $q + \epsilon < s$ about the origin as center and ϵ is a positive constant such that $q + \epsilon < f$.

2.2. Properties of series of functions $A_{n,r}(x)$. Suppose we have a series

$$F(x) = \sum_{n=0}^{\infty} c_n A_{n,r}(x), \tag{2.3}$$

where r is a fixed non-negative integer, satisfying the conditions of Theorem 1.1 as to convergence. We seek to determine the class of functions $F(x)$ represented by such convergent series. Let us set

$$\limsup_{n=\infty} |\, c_n \,|^{1/n} = q. \tag{2.4}$$

Then, as we have seen, $q \leqq \gamma_{r+1}$ and the function $\phi(z)$ defined by (1.11) is analytic for $|z| > q$. Let us form the function

$$H(x) = \frac{1}{2\pi i}\int_{C_{r+1}} \frac{E(x,z)}{g(z)}\phi(z)\,dz \tag{2.5}$$

which by Lemma 2.1 is of sort $\{u_n\}$, type not exceeding γ_{r+1}. Integrating term by term the expression obtained by substituting the expansion (1.11) of $\phi(z)$ into (2.5), we see that

$$H(x) = \sum_{n=0}^{\infty} c_n A_{n,r+1}(x). \tag{2.6}$$

The difference $H(x) - F(x)$ is given by the expression (1.9). By Corollary 1.1 the convergence of the series (2.3) implies the convergence of each of the series (1.10) and hence we may separate the sum in (1.9) into a finite number of terms of the sort

$$D_m{}^k E(x,z)\sum_{n=0}^{\infty}\frac{c_n}{\gamma_{r+1}^{n+1}}\binom{n+k}{k}e^{-i(n+1)\theta_m}.$$

From the form of $D_m{}^k E(x,z)$ we see that it is of sort $\{u_n\}$, type $|\alpha_m|$. Thus the function $F(x)$, being expressible as a finite sum of functions each of sort $\{u_n\}$, type not exceeding γ_{r+1}, is itself of sort $\{u_n\}$, type not exceeding γ_{r+1}.

THEOREM 2.1. *If a function $F(x)$ has the convergent expansion (2.3) then $F(x)$ is of sort $\{u_n\}$, type not exceeding γ_{r+1}.*

Expressing the series in (2.3) as a series of the form (2.1) and using property 2° of the sequence $\{u_n(x)\}$, we have the following theorem:

THEOREM 2.2. *No function $F(x)$ can have two distinct expansions in series of the form (2.3) for the same value of r.*

Let $F(x)$ be a given function of sort $\{u_n\}$, type q, where $\gamma_r \leqq q < \gamma_{r+1}$. We study the problem of expanding $F(x)$ in a series of the form (2.3). The function $F(x)$ does have an expansion of the form (2.1) with (2.2) holding and hence we may write

$$F(x) = \frac{1}{2\pi i}\int_{C_r} \frac{E(x,z)}{g(z)}\,g(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz, \tag{2.7}$$

where $C'_r = q + \epsilon < \gamma_{r+1}$ and ϵ is positive. Writing

$$g(z)\sum_{n=0}^{\infty} e_n z^{-n-1} = \sum_{n=-\infty}^{\infty} c_n z^{-n-1} = \sum_{n=0}^{\infty} c_n z^{-n-1} + P(z), \qquad q < |z| < \rho,$$

we have

$$(2.8) \qquad F(x) = \frac{1}{2\pi i}\int_{C_r} \frac{E(x,z)}{g(z)}\,[P(z) + \sum_{n=0}^{\infty} c_n z^{-n-1}]\,dz.$$

Integrating term by term, we have (see (1.2))

$$(2.9) \qquad F(x) = \sum_{m=1}^{\mu(r)} \sum_{k=0}^{s_m-1} d_{mk} D_m{}^k E(x,z) + \sum_{n=0}^{\infty} c_n A_{n,r}(x),$$

where the c's and d's are uniquely determined constants. If $r = 0$, the finite sum on the right of (2.8) vanishes since C_0 contains no zero of $g(z)$.

THEOREM 2.3. *Any given function $F(x)$ of sort $\{u_n\}$, type q, such that $\gamma_r \leqq q < \gamma_{r+1}$ has an expansion of the form (2.9) where the c's and d's are uniquely determined constants.*

We may also write $F(x)$ in the form (2.8) with C_r replaced by C_{r+p}, where p is any positive integer. Since the Laurent's series in (2.8) is valid for z in an annular ring which includes the circle C_{r+p} we have the following corollary:

COROLLARY 2.1. *The function $F(x)$ has an expansion in the form (2.9) with r replaced throughout by $r + p$, where p is any positive integer. The constants c_n $(n = 0, 1, \cdots)$, and d_{mk} $(k = 0, \cdots, s_m - 1;\ m = 1, \cdots, \mu(r))$, are the same in every case.*

Suppose $F(x)$ is a given function of sort $\{u_n\}$, type γ_{r+1}. Then, by Theorem 2.3, it has an expansion of the form

$$(2.10) \qquad F(x) = \sum_{m=1}^{\mu(r+1)} \sum_{k=0}^{s_m-1} d_{mk} D_m{}^k E(x,z) + \sum_{n=0}^{\infty} c_n A_{n,r+1}(x).$$

If $F(x)$ also has an expansion of the form (2.9) it is easily seen that the coefficients c_n be the same in the two cases. Since the series (1.9) represents the difference $\Sigma c_n A_{n,r+1}(x) - \Sigma c_n A_{n,r}(x)$ the expansions (2.9) and (2.10) can coexist only if the series (1.9) converges. In section 1.2 we proved that the convergence of the series (1.9) implies the convergence of each of the series (1.10). From its form we see that if (1.9) converges it gives the negative of the terms for $m = \mu(r) + 1, \cdots, \mu(r+1)$ in the finite sum in (2.10); furthermore, when the series (1.9) converges the two expansions (2.9) and (2.10) coexist.

This completes the proof of the following theorem:

THEOREM 2.4. *Let $F(x) = \Sigma e_n u_n(x)$ be any given function of sort $\{u_n\}$, type γ_{r+1}. Write*

$$\frac{1}{2\pi i}\int_{C_{r+1}} E(x,z)g(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz = \sum_{n=0}^{\infty} c_n u_n(x).$$

Then a necessary and sufficient condition that $F(x)$ have an expansion of the form (2.9) *is that each of the series* (1.10) *shall converge.*

These theorems give necessary and sufficient conditions for the expansion of functions in series of the functions $A_{n,r}(x)$.

III. Functional Equations.

3.1. The operator B. We define the operator B to be such that if $F(x)$ be any given function of sort u_n, type less than f, that is, if

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x), \qquad \limsup_{n=\infty} |\, e_n \,|^{1/n} < f, \tag{3.1}$$

then

$$B^m F(x) = \sum_{n=0}^{\infty} e_{n+m} u_n(x) \qquad (m = 0, 1, \cdots). \tag{3.2}$$

Expanding $A_{n,r}(x)$ $(r = 0, 1, \cdots;\ n = 0, \pm 1, \pm 2, \cdots)$, in series of the form (3.1), we see that $B\,A_{n,r}(x) = A_{n-1,r}(x)$.

3.2. Existence theorems. Let $h(z) = \sum_0^{\infty} b_n z^n$, $b_0 \neq 0$, be a series convergent for $|\, z \,| < \sigma$ where $0 < \sigma \leqq f$. Let q be any positive number less than σ. Let $\theta(x) = \sum_0^{\infty} c_n u_n(x)$ be any given function of sort $\{u_n\}$, type not exceeding q. Let us seek solutions

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x) \tag{3.3}$$

of the linear functional equation

$$b_0 F(x) + b_1 BF(x) + b_2 B^2 F(x) + \cdots = \theta(x), \tag{3.4}$$

under the following hypothesis on the coefficients:

$$\limsup_{n=\infty} |\, e_n \,|^{1/n} \leqq q. \tag{3.5}$$

The trial solution (3.3) may be written in the form

$$F(x) = \frac{1}{2\pi i}\int_C E(x,z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz, \tag{3.6}$$

where C is any circle about the origin as center of radius $q + \epsilon < \sigma$ and ϵ is a positive number such that $h(z)$ has no zero in the ring $q < |\, z \,| \leqq q + \epsilon$.

Substituting the trial solution (3. 6) into the equation (3. 4) and performing an interchange of operations, we see that a necessary and sufficient condition that (3. 3) or its equivalent (3. 6) be a solution of (3. 4) is that

$$\theta(x) = \frac{1}{2\pi i}\int_C E(x,z)h(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz, \tag{3.7}$$

which holds if and only if the e's satisfy the conditions

$$\sum_{\mu=0}^{\infty} b_\mu e_{\mu+n} = c_n \qquad (n = 0, 1, \cdots). \tag{3.8}$$

A particular solution of (3. 4) of sort $\{u_n\}$, type not exceeding q is obviously given by

$$F_1(x) = \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\sum_{n=0}^{\infty} c_n z^{-n-1}dz, \tag{3.9}$$

and hence a set of solutions $\{e_n\}$ of (3. 8) of the character described by (3. 5) is given by setting $e_n = d_n$ $(n = 0, 1, \cdots)$, where the d_n are defined by

$$[h(z)]^{-1}\sum_{n=0}^{\infty} c_n z^{-n-1} = \sum_{n=0}^{\infty} d_n z^{-n-1} + \sum_{n=0}^{\infty} f_n z^n, \qquad q < |z| \leqq q + \epsilon.$$

If we write the trial solution (3. 3) (or (3. 6)) in the form

$$\begin{aligned} F(x) &= \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\, h(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz \\ &= \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\cdot\left[\sum_{n=0}^{\infty}\delta_n z^{-n-1} + \sum_{n=0}^{\infty}\epsilon_n z^n\right]dz, \end{aligned}$$

then we see that $F(x)$ is a solution of (3. 4) if and only if $\delta_n = c_n$. In such a case

$$F(x) = F_1(x) + \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\sum_{n=0}^{\infty}\epsilon_n z^n dz.$$

If $h(z)$ has no zero for $|z| \leqq q$, then $F_1(x)$ is the only solution of (3. 4) of sort $\{u_n\}$, type not exceeding q. If $h(z)$ has zeros of multiplicities σ_m at points

$$a_m, \qquad |a_m| \leqq q \qquad (m = 1, \cdots, k), \tag{3.10}$$

then by the theory of residues it follows that $F(x)$ must be of the form

$$F(x) = F_1(x) + \sum_{m=1}^{k}\sum_{j=0}^{\sigma_m - 1}\frac{\partial^j}{\partial z^j}E(x,z)|_{z=a_m}. \tag{3.11}$$

Every solution of (3. 4) of sort $\{u_n\}$, type not exceeding q, is expressible in the form (3. 11) where the d's are constants. But each of the functions

$$(3.12)\qquad \frac{\partial^j}{\partial z^j} E(x,z)|_{z=a_m} = \sum_{n=j}^{\infty} n(n-1)\cdots(n-j+1)a_m^{n-j}u_n(x), \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k),$$

constitutes a solution of the homogeneous equation, that is, of equation (3.4) formed for $\theta(x) \equiv 0$. These solutions are linearly independent (see Lemma 1.1) and each is of sort $\{u_n\}$, type not exceeding q.

This completes the proof of the following theorem:

THEOREM 3.1. *Let* $h(z) = \sum_0^\infty b_n z^n$, $b_0 \neq 0$, *be a series convergent for* $|z| < \sigma \leqq f$. *Let* q *be any positive number less than* σ. *Let* $\theta(x) = \sum_0^\infty c_n u_n(x)$ *be any function of sort* $\{u_n\}$, *type not exceeding* q. *If* $h(z)$ *has no zero for* $|z| \leqq q$, *then the only solution of the equation* (3.4) *of sort* $\{u_n\}$, *type not exceeding* q, *is given by* (3.9). *If* $h(z)$ *has zeros of multiplicities* σ_m *at the points* (3.10) *then the most general solution of* (3.4) *of sort* $\{u_n\}$, *type not exceeding* q, *is given by* (3.11) *where the d's are arbitrary constants.*[8]

For the case where $h(z) = g(z)$ the solutions $F(x)$ are expressible in the form (2.9) where r is the integer defined by $\gamma_r \leqq q < \gamma_{r+1}$.

Noting that every solution $F(x)$ of (3.4) of the desired character, when expressed in the form (3.3), furnishes a set of solutions $\{e_n\}$ of (3.8), verifying (3.5), we obtain as corollaries certain results relating to infinite systems of linear equations in an infinite number of unknowns. We state only the following one relating to homogeneous systems:[9]

COROLLARY 3.1. *Let* $h(z) = \sum_0^\infty b_n z^n$, $b_0 \neq 0$, *be a series convergent for* $|z| < \sigma$. *Let* q *be any positive number less than* σ. *If* $h(z)$ *has no zero for* $|z| \leqq q$, *then the only set of solutions* $\{e_n\}$ *of*

$$(3.13)\qquad \sum_{\mu=0}^{\infty} b_\mu e_{\mu+n} = 0, \qquad (n=0,1,\cdots),$$

verifying (3.5), *is the identically zero one. If* $h(z)$ *has* p *zeros (multiple zeros counted multiply),* $p > 0$, *for* $|z| \leqq q$, *then there exist exactly* p *linearly independent sets of solutions* $\{e_n\}$ of (3.13) *of the character described by*

[8] Since $\theta(x)$ may be identically zero the theorem applies also to homogeneous equations.

[9] Perron (*loc. cit.*) by different methods has obtained the results contained in this corollary.

(3.5). *If these p zeros are given by* (3.10), *where* a_m *is of multiplicity* σ_m, *then p such sets of solutions are given by*

$$e_n = n(n-1)\cdots(n-j+1)a_m^{n-j} \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k), \tag{3.14}$$

or by a linear combination of these, namely

$$e_n = n^j a_n{}^m \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k). \tag{3.15}$$

3.3. A particular case of the operator B. Let $\lambda_0, \lambda_1, \lambda_2, \cdots$, be an infinite sequence of non-zero constants such that

$$\lim_{n=\infty} |\lambda_0\lambda_1\cdots\lambda_n|^{1/n} = \infty, \qquad \lim_{n=\infty} |\lambda_n|^{1/n} = 1. \tag{3.16}$$

The sequence $\{u_n(x) = x^n/(\lambda_0\lambda_1\cdots\lambda_n)\}$ forms a sequence possessing the two properties 1^0 and 2^0 in the introduction. We give here a definition of the operator B associated with this sequence which is an extension of the one already given. If $F(x)$ be any single-valued analytic function regular at $x=0$ then we define

$$B^mF(x) = \frac{1}{2\pi i}\int_C \frac{F(z)}{z^{m+1}} \sum_{n=0}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n+m}}{\lambda_0\lambda_1\cdots\lambda_n}\frac{x^n}{z^n}\,dz \quad (m=0,1,\cdots), \tag{3.17}$$

where C is any circle about the origin as center within the region of analyticity of $F(z)$ and x is any interior point of C. In view of (3.16) each of these integrals is defined for any point x interior to C and the Maclaurin's expansion of $B^mF(x)$ has the same circle of convergence as has the one representing $F(x)$. We note the following additional properties of the operator B:

$$Bx^n = \lambda_n x^{n-1}, \quad n>0; \qquad Bc = 0,\ c \text{ any constant.}$$

If

$$F(x) = \sum_{n=0}^{\infty} \delta_n \frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n}, \qquad |x|<R,$$

then

$$B^mF(x) = \sum_{n=0}^{\infty} \delta_{n+m} \frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n}, \qquad |x|<R.^{10} \tag{3.18}$$

[10] Pincherle and Hadamard have studied operators of this sort. See S. Pincherle, *Giornale di Matematiche*, vol. 32 (1884), pp. 62-74; J. Hadamard, *Acta Mathematica*, vol. 22 (1899), pp. 55-63. Hadamard introduced and used a contour integral representation of the form (3.17) in his study of the *Hadamard product*. Carmichael suggested to me the above contour integral representation for the operation.

We may and will also consider (3.17) and (3.18) defined for m a negative integer provided that we agree to replace by zero any term which contains a λ or a δ with a negative subscript.

For the case $\lambda_0 = 1$, $\lambda_n = n$, $n > 0$, the operation $BF(x)$ furnishes the derivative of an analytic function regular at $x = 0$. For negative values of m, say $m = -k$, equations (3.17) and (3.18) represent generalizations of the multiple integrals

$$\int_0^x \int_0^{z_1} \cdots \int_0^{z_{k-1}} F(z_k)\, dz_k \cdots dz_2 dz_1.$$

We place a further restriction on the λ's, namely that

$$|\lambda_0| \leqq |\lambda_1| \leqq |\lambda_2| \leqq \cdots. \tag{3.19}$$

Carmichael [11] has obtained a general expansion theorem relating to a set of non-zero constants verifying (3.19) and such that $\lim |\lambda_n| = \infty$. I state it here for future reference.

GENERAL EXPANSION THEOREM. *Let $a_1(x), a_2(x), \cdots$, be a finite or infinite sequence of functions each of which is analytic in the interior of the circle $|x| = \bar{\rho}$. Let a be a positive number less than $\bar{\rho}$. Denote by M_k the maximum absolute value of $a_k(x)$ in the region $|x| \leqq a$. Suppose the following series converges*

$$M_1 \frac{a}{|\lambda_1|} + M_2 \frac{a^2}{|\lambda_1\lambda_2|} + M_3 \frac{a^3}{|\lambda_1\lambda_2\lambda_3|} + \cdots. \tag{3.20}$$

Let $F(x)$ be analytic for $|x| \leqq a$. Then $F(x)$ has a unique expansion in either of the following two forms:

$$F(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{\lambda_0\lambda_1 \cdots \lambda_n} \left\{ 1 + a_1(x) \frac{x}{\lambda_{n+1}} + a_2(x) \frac{x^2}{\lambda_{n+1}\lambda_{n+2}} + a_3(x) \frac{x^3}{\lambda_{n+1} \cdots \lambda_{n+3}} + \cdots \right\}, \tag{3.21}$$

$$F(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{\lambda_0\lambda_1 \cdots \lambda_n} + a_1(x) \sum_{n=0}^{\infty} c_n \frac{x^{n+1}}{\lambda_0\lambda_1 \cdots \lambda_{n+1}} + a_2(x) \sum_{n=0}^{\infty} c_n \frac{x^{n+2}}{\lambda_0\lambda_1 \cdots \lambda_{n+2}} + \cdots. \tag{3.22}$$

[11] R. D. Carmichael, "A general expansion theorem with applications to certain integral equations of infinite order." An abstract of this paper is given in *Bulletin of the American Mathematical Society*, vol. 40 (1934), p. 211. The theorem, although actually stated for the case of positive λ's, carries with it the case of complex λ's, under the above conditions.

Either of the expansions of $F(x)$ converges absolutely and uniformly in any whatever circle interior to the circle $|x| = a$, and is valid for every interior point of the latter circle.

In the following paragraph we will use the operator B defined as in equation (3.17) where the λ's satisfy the additional conditions (3.19). By a method analogous to that used by Carmichael (*op. cit.*) we obtain theorems with reference to the operator B similar to those obtained by Carmichael with reference to the differential operator. Because of their similarity as to proof and statement we merely state one.

THEOREM 3.3. *Under the hypotheses on $F(x)$, $a_1(x)$, $a_2(x), \cdots$, stated in the General Expansion Theorem, the mixed equation*

$$(3.23)\quad F(x) = B^n v(x) + a_1(x)B^{n-1}v(x) + \cdots + a_n(x)v(x) + a_{n+1}(x)B^{-1}v(x) + a_{n+2}(x)B^{-2}v(x) + \cdots,$$

where n is a positive integer, has one and just one solution $v(x)$ which is analytic at $x = 0$ and which satisfies the initial conditions $v(0) = \zeta_0$, $Bv(0) = \zeta_1, \cdots, B^{n-1}v(0) = \zeta_{n-1}$ where $\zeta_0, \zeta_1, \cdots, \zeta_{n-1}$ are any preassigned constants. This solution is analytic in the interior of the circle $|x| = a$.

UNIVERSITY OF ILLINOIS.

ON THE BOUNDARY OF THE RANGE OF VALUES OF $\zeta(s)$.

By RICHARD KERSHNER and AUREL WINTNER.

In connection with his investigation of the distribution of the values of the Riemann zeta function,[1] Bohr[2] has studied the vectorial addition[3] of convex curves and has proven, among other things, that if the vectorial sum of a finite or infinite sequence of convex curves is a bounded set, then it is either a closed convex region or a ring-shaped region bounded by two non-intersecting convex curves. The method of Bohr is entirely geometrical and it is, therefore, impossible to apply it to a discussion of smoothness properties of the boundary of the vectorial sum, at least if the sum is infinite. Following a suggestion of one of the present authors, Haviland[4] applied the supporting function (Stützfunktion) of Brunn and Minkowski to the study of Bohr's problem, using the fact that the supporting function of the outer boundary of the vectorial sum is the sum of the supporting functions of the added curves. Bohr and Jessen[5] have recently shown that, with the use of the supporting functions, it is possible to determine the set of those $\sigma_0 > 1$ for which the closure of the values attained by the logarithm of the Riemann zeta function on the line $\sigma = \sigma_0$ is a convex region; for the remaining values of $\sigma_0 > 1$ this closure is ring-shaped. It has been shown by one of the present authors[6] that a rule similar to the above mentioned rule for the supporting function of the outer boundary holds for the supporting function of the inner boundary, at least on arcs of the inner boundary which are free of corners.

In the present paper the method of supporting functions will be applied to the question of regular analyticity of the boundary curves of the infinite vectorial sum in cases of the type of the logarithm of the Riemann zeta function. In particular, it will be shown that the outer boundary of the closure

[1] Cf. E. C. Titchmarsh, *The Zeta Function of Riemann*, Cambridge, 1930, Chapter IV.

[2] H. Bohr, "Om Addition af uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab*, Forhandlinger, 1913, pp. 325-366.

[3] By the vectorial sum $A_1 + A_2$ of two sets A_1 and A_2 is meant the set of all points $z = z_1 + z_2$, where $z_1 \subset A_1$ and $z_2 \subset A_2$. It is clear that this addition is associative and commutative. By an infinite vectorial sum $A_1 + A_2 + \cdots$ is meant the set of points z which may be expressed in at least one way as a limit of points $z^{(n)} = z_1 + z_2 + \cdots z_n$, where $z_k \subset A_k$.

[4] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[5] H. Bohr and B. Jessen, "On the distribution of the values of the Riemann zeta function," *American Journal of Mathematics*, vol. 58 (1936), pp. 35-44.

[6] R. Kershner, "On the addition of convex curves." To appear later. For an abstract, cf. *Bulletin of the American Mathematical Society*, vol. 42 (1936). Record of the February, 1936, meeting.

of the values attained by $\log \zeta(s)$ on a fixed line $\sigma = \sigma_0 > 1$ is a regular analytic curve. Due to the explicit formula for the supporting function of the inner curve, the analyticity of the inner curve may be treated similarly.[6] The result is of particular interest in view of the fact that the density of the asymptotic distribution cannot be analytic at points of the boundary. The transition from the range of $\log \zeta(\sigma_0 + it)$ to that of $\zeta(\sigma_0 + it)$ requires but a trivial exponential mapping.

It is known that if the power series

$$p(w) = a_1 w + a_2 w^2 + \cdots, \quad \text{where} \quad a_1 \neq 0, \tag{1}$$

is convergent in the vicinity of $w = 0$, then there exists an $R > 0$ such that $p(w)$ is regular and schlicht in $|w| \leqq R$ and that[7] the curve $z = p(re^{i\theta})$, where $0 \leqq \theta < 2\pi$, is, for every fixed positive value of $r \leqq R$, a regular analytic convex Jordan curve in the (x, y)-plane, where $z = x + iy$. Let $r_1, r_2, \cdots$ be an infinite sequence of positive numbers such that $r_n < R$ for every n and that $r_1 + r_2 + \cdots$ is convergent. Let C_n denote the convex curve

$$z = z_n(\theta) = p(r_n e^{i\theta}) \tag{2}$$

and $T_n(\theta)$ the tangent to C_n at the point $z = p(r_n e^{i\theta})$. Then the equation of $T_n(\theta)$ is

$$\xi \cos\phi + \eta \sin\phi - h_n(\phi) = 0, \tag{3}$$

where

$$\cos\phi = -\dot{y}_n(\theta)/|\dot{z}_n(\theta)|, \qquad \sin\phi = \dot{x}_n(\theta)/|\dot{z}_n(\theta)| \tag{4}$$

and

$$h_n(\phi) = x_n(\theta)\cos\phi + y_n(\theta)\sin\phi, \tag{5}$$

it being understood that $z_n = x_n + iy_n$ and that the dot denotes differentiation with respect to θ. Since

$$e^{i\phi} = \{-\dot{y}_n(\theta) + i\dot{x}_n(\theta)\}/|\dot{z}_n(\theta)| = i\dot{z}_n(\theta)/|\dot{z}_n(\theta)|,$$

it is clear from (2) that

$$e^{i\phi} = -e^{i\theta} p'(r_n e^{i\theta})/|p'(r_n e^{i\theta})|, \tag{6}$$

where the prime denotes differentiation of $p(w)$ with respect to w. On placing $\bar{p}(w) = \bar{a}_1 w + \bar{a}_2 w^2 + \cdots$, it follows from (6) that

$$e^{i\phi} = -e^{i\theta} p'(r_n e^{i\theta})/\{p'(r_n e^{i\theta})\bar{p}'(r_n e^{-i\theta})\}^{1/2},$$

so that

$$\phi = \phi_n(\theta) = \pi + \theta - i\log p'(r_n e^{i\theta}) + (i/2)\log p'(r_n e^{i\theta}) + (i/2)\log \bar{p}'(r_n e^{-i\theta}). \tag{7}$$

---

[7] Cf., e. g., G. Pólya und G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1, Berlin, 1925, p. 105, no. 108.

Hence

$$\frac{d\phi_n(\theta)}{d\theta} = 1 + \frac{r_n e^{i\theta} p''(r_n e^{i\theta})}{p'(r_n e^{i\theta})} - \frac{r_n e^{i\theta} p''(r_n e^{i\theta})}{2p'(r_n e^{i\theta})} + \frac{r_n e^{-i\theta} \bar{p}''(r_n e^{-i\theta})}{2\bar{p}'(r_n e^{-i\theta})}. \tag{8}$$

Now $p(w)$ and, consequently, $\bar{p}(w)$ are regular and schlicht in $|w| \leqq R$, which implies that the absolute values of their first derivatives have a positive lower bound β in this circle. The absolute values of the second derivatives of $p(w)$ and $\bar{p}(w)$ have in this circle a finite upper bound γ. Consequently, since $r_n \leqq R$, it is seen from (8) that

$$\left| \frac{d\phi_n(\theta)}{d\theta} \right| \geqq 1 - \alpha r_n,$$

where the positive number $\alpha = 2\gamma/\beta$ is independent of both n and θ. Thus, since $r_n \to 0$, there exists a sufficiently large n_0 such that

$$\left| \frac{d\phi_n(\theta)}{d\theta} \right| \geqq \tfrac{1}{2} \quad \text{whenever} \quad n \geqq n_0. \tag{9}$$

Now the equation (6) may be considered as defining θ as a univalued, continuous function

$$\theta = \theta_n(\phi) \tag{10}$$

of the normal inclination ϕ of $T_n(\theta)$. In fact, since C_n is a regular analytic convex Jordan curve, the two angular parameters θ and ϕ are in a continuous one-to-one correspondence, so that (10) is the inverse function of the function (7). Finally, it is seen from (3) and (5) that the supporting function of C_n is

$$h_n(\phi) = x_n(\theta_n(\phi)) \cos\phi + y_n(\theta_n(\phi)) \sin\phi. \tag{11}$$

Let $\rho > 0$ be so small that the functions (2), where $n = 1, 2, \cdots$, are regular and uniformly bounded functions of the *complex* variable θ in the rectangle

$$-\rho < \Re\theta < 2\pi + \rho, \qquad -\rho < \Im\theta < \rho. \tag{12}$$

The existence of such a $\rho > 0$, which is independent of n, is obvious from (2), since, on the one hand, every r_n is chosen less than the convergence radius of (1) and, on the other hand, $r_n \to 0$ in view of the convergence of $r_1 + r_2 + \cdots$. Since $p'(w)$ and $\bar{p}'(w)$ do not vanish in $|w| \leqq R$, one may choose n_0 so large that not only does (9) hold but also the functions $\phi = \phi_n(\theta)$, when defined in the *complex* region (12) by the explicit formula (7), are regular and in absolute value less than a number M which is independent of $n(\geqq n_0)$.

Let $\theta = \theta_0$ be a fixed real angle in the interval $0 \leqq \theta_0 \leqq 2\pi$ and let ϕ_0 be the corresponding angle ϕ. Then, in the circle $|\theta - \theta_0| < \rho$, the functions $\phi = \phi_n(\theta)$, where $\phi_0 = \phi_n(\theta_0)$, are regular analytic and in absolute value less

than M for every $n \geqq n_0$. Furthermore, $\left| \frac{d\phi_n(\theta)}{d\theta} \right|_{\theta=\theta_0} \geqq \frac{1}{2}$ by (9). Hence, by the proof of the local theorem on the inverse function of an analytic function, there exists [8] a positive constant τ which is independent of θ_0 and of n ($\geqq n_0$) and is such that the inverse function $\theta = \theta_n(\phi, \theta_0, \phi_0)$, where $\phi_0 = \phi_n(\theta_0)$, is regular analytic and $| \theta - \theta_0 | < \rho$ in the circle $| \phi - \phi_0 | < \tau$. Now it is obvious from the monodromy theorem that $\theta_n(\phi, \theta_0, \phi_0)$ is independent of θ_0 and of $\phi_0 = \phi_n(\theta_0)$, so that it may be denoted simply by $\theta_n(\phi)$. Thus, if $n \geqq n_0$, all functions $\theta_n(\phi)$ are regular analytic and $\theta = \theta_n(\phi)$ is within the rectangle (12), hence the functions $\theta_{n_0}(\phi), \theta_{n_0+1}(\phi), \cdots$ are uniformly bounded, if ϕ is in the rectangle

$$(13) \qquad \mu - \tau/2 \leqq \Re\phi \leqq \mu + 2\pi + \tau/2, \qquad -\tau/2 < \Im\phi < \tau/2,$$

where $\mu = \phi_n(0)$ is the normal inclination of $T_n(0)$. Consequently, the functions $\theta_n(\phi)$, which were defined for $\mu \leqq \phi_0 \leqq \mu + 2\pi$ by (10), are regular and uniformly bounded in the rectangle (13) of the complex ϕ-plane for every $n \geqq n_0$. Finally, since the intervals $\mu \leqq \phi_0 \leqq \mu + 2\pi$ and $0 \leqq \theta_0 \leqq 2\pi$ are in a one-to-one correspondence, and since the derivative $d\phi/d\theta$ is, for real θ, distinct from zero for every n, it is clear from the definition (10) of $\theta_n(\phi)$ that there exists, for every fixed n, a rectangle R_n in the complex ϕ-plane such that R_n contains the interval $\mu \leqq \phi_0 \leqq \mu + 2\pi$ in its interior and the function (10) is regular and bounded in R_n. Consequently, there exists in the complex ϕ-plane a rectangle R_0 in which *all* functions $\theta_n(\phi)$ are regular and uniformly bounded and which contains the interval $\mu \leqq \phi_0 \leqq \mu + 2\pi$ in its interior. In fact, if Q denotes the rectangle (13), the common part of the finite number of rectangles Q; $R_1, R_2, \cdots, R_{n_0-1}$ will be such an R_0.

Now it will be shown that

$$(14) \qquad H(\phi) = \sum_{n=1}^{\infty} h_n(\phi)$$

is regular analytic in the interval $\mu \leqq \phi \leqq \mu + 2\pi$, i. e., in an open rectangle of the complex ϕ-plane which contains this interval. It has been shown above that if ϕ is in the rectangle R_0, then $| \theta_n(\phi) | < L$ for some positive constant L. On the other hand, it is clear from (1) and (2), where $z_n = x_n + iy_n$, that, since $r_n \to 0$, one may choose a sufficiently large $K = K_L$ and then a positive constant $A = A_K$ such that, if $| \theta | < L$ and $n \geqq K$, the functions $x_n(\theta)$ and $y_n(\theta)$ are regular and in absolute value not larger than Ar_n. Consequently, $x_n(\theta_n(\phi))$ and $y_n(\theta_n(\phi))$ are regular and in absolute value not larger than

[8] Cf. e. g., E. Landau, "Der Picard-Schottkysche Satz und die Blochsche Konstante," *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalische-mathematische Klasse*, 1926, p. 470.

Ar_n if ϕ is in the rectangle R_0 and $n \geqq K$. Hence, since $r_1 + r_2 + \cdots$ is convergent by assumption, the series

$$x(\phi) = \sum_{n=1}^{\infty} x_n(\theta_n(\phi)), \qquad y(\phi) = \sum_{n=1}^{\infty} y_n(\theta_n(\phi))$$

have regular analytic terms and are uniformly convergent in the rectangle R_0 of the ϕ-plane. It follows, therefore, from (11) that the function (14) is regular in R_0. This completes the proof of the analyticity of (14) in a rectangle which contains the interval $\mu \leqq \phi \leqq \mu + 2\pi$ in its interior.

Now $H(\phi)$ is [9] the supporting function of the outer curve. On introducing into the regular analytic functions $x = x(\phi)$, $y = y(\phi)$ the new independent variable

$$s = \int^{\phi} [(x'(\psi))^2 + (y'(\psi))^2]^{1/2}\, d\psi$$

instead of ϕ, it follows that the outer boundary of the region $C_1 + C_2 + \cdots$ is a regular analytic curve.

If $p(w)$ is the power series

$$p(w) = -\log(1 - w), \quad \text{where} \quad p(0) = 0,$$

then $p(w)$ is regular and schlicht for $|w| < 1$ and the image of the circle $|w| = r$ is convex for every $r < 1$. Now it is known [10] that the closure of the values of the logarithm of the Riemann zeta function, $\log \zeta(s)$, on a fixed line $\Re s = \sigma > 1$, is the region $C_1 + C_2 + \cdots$, where C_n is the curve

$$z = -\log(1 - p_n^{-\sigma} e^{i\theta})$$

and p_n is the n-th prime number. Since $p_n^{-\sigma} < 1$ for every n and $p_1^{-\sigma} + p_2^{-\sigma} + \cdots$ is convergent for $\sigma > 1$, it follows that the outer boundary of the closure of the values of $\log \zeta(s)$ on a fixed line $\sigma > 1$ is a regular analytic convex curve. The inner curve, if any, is similarly treated up to its possible corners.

THE JOHNS HOPKINS UNIVERSITY.

ERRATUM.

Professor H. Cramér has kindly pointed out to me that Theorem XVIII of my paper "On a class of Fourier transforms" (pp. 45-90 of vol. 58 (1936) of this JOURNAL) is false unless one reads at the end of this Theorem "a Gauss-Maxwell law" instead of "the Gauss-Maxwell law of radial symmetry." This change is necessitated by the fact that on p. 81 the statement "the matrix $\| -\frac{1}{2}\mu_{pq} \|$ is $C(e_u)$ times the unit matrix" is erroneous. That Theorem XVIII becomes correct by omitting the restriction of radial symmetry is seen from the central limit theorem.

A. WINTNER.

---

[9] Cf. E. K. Haviland, *loc. cit.*

[10] Cf., e. g., E. C. Titchmarsh, *loc. cit.*, pp. 64-67.

THE SAMPLING THEORY OF SYSTEMS OF VARIANCES, COVARIANCES AND INTRACLASS COVARIANCES.

By S. S. WILKS.

It is well known that the type of data to which the method of intraclass correlation is applicable can be effectively and accurately treated by the method of the Analysis of Variance. The connection between the two methods has been given by Fisher [1] for the case of one variable, and it is quite evident that the latter method has the greater practical value in determining, statistically, the relative importance of two groups of factors causing variation. However, the method of intraclass correlation has theoretical value in that it provides estimates of the parameters which characterize the distribution of the variates which constitute a "family." For example, under the assumption of normality, if the mean and variance of each of the variates $x_1, x_2, \cdots x_k$ is m and σ^2 and the coefficient of correlation between any pair is ρ, then the distribution of this k-variate "family" is the normal multivariate distribution characterized, however, by only the parameters m, σ^2 and ρ. ρ is called the intraclass correlation coefficient for the k variates. The method of intraclass correlation consists essentially in estimating m, σ^2 and ρ and making a significance test appropriate to the hypothesis under consideration.

In the present note we shall demonstrate a general method for handling the sampling theory of systems of variances, covariances and intraclass covariances, and illustrate the method on the generalized intraclass correlation problem.

In order to present the essentials of the method, we shall consider the following general problem. Let a system of variates $y_p = x_p - m_p$ $(p = 1, 2, \cdots n)$ be distributed according to the law

$$(1) \qquad U_n(A_{pq}, y_p) = \frac{|A_{pq}|^{1/2}}{(\pi)^{n/2}} \exp(-\Sigma A_{pq} y_p y_q) \prod_{p=1}^{n} dy_p$$

where the matrix $\| A_{pq} \|$ is positive definite with determinant $|A_{pq}|$. The characteristic function of the quantities $\xi_{pq} = y_p y_q$ is given by

$$(2) \qquad \begin{aligned} \phi = E(\exp(i\Sigma t_{pq}\xi_{pq})) &= \int \exp(i\Sigma t_{pq}\xi_{pq}) \cdot U_n(A_{pq}, y_p) \\ &= |A_{pq}|^{1/2} \, |A_{pq} - it_{pq}|^{-(1/2)} \end{aligned}$$

[1] R. A. Fisher, *Statistical Methods for Research Workers*, 4th edition (1932), p. 205. Also see R. A. Fisher, *Metron*, vol. 1, no. 4 (1921), pp. 3-32.

where $t_{pq} = t_{qp}$ and the integral is taken over all values of the y_p. Now suppose ϕ factors into determinants of order m as follows

$$\phi = \prod_{\alpha=1}^{g} | B^{\alpha}_{uv} |^{k\alpha/2} | \overline{B^{\alpha}}_{uv} |^{-(k\alpha/2)} \tag{3}$$

where the k_α are positive integers, $\| B^{\alpha}_{uv} \|$ is positive definite and

$$B^{\alpha}_{uv} = \sum_{\beta=1}^{g} c_{\alpha\beta} A^{\beta}_{uv}, \qquad \bar{B}^{\alpha}_{uv} = B^{\alpha}_{uv} - i \sum_{\beta=1}^{g} c_{\alpha\beta} t^{\beta}_{uv}, \qquad | c_{\alpha\beta} | \neq 0 \tag{4}$$

the A^{α}_{uv} (where $A^{\alpha}_{uv} = A^{\alpha}_{vu}$) being elements of the matrix $\| A_{pq} \|$. The corresponding elements in $\| t_{pq} \|$ will be denoted by t^{α}_{uv}. $\{A^{\alpha}_{uv}\}$ and $\{t^{\alpha}_{uv}\}$ will be the sets of distinct elements in $\| A_{pq} \|$ and $\| t_{pq} \|$ respectively. Therefore, since the ξ_{pq} are all distinct, we have

$$\sum_{p,q} t_{pq} \xi_{pq} = \sum_{u,v,\beta} t^{\beta}_{uv} \eta^{\beta}_{uv} \tag{5}$$

where $\eta^{\beta}_{uv} = \Sigma \xi_{pq}$, summed over all values of p and q for which $t_{pq} = t^{\beta}_{uv}$.

Now if we let $\sum_{\beta=1}^{g} c_{\alpha\beta} t^{\beta}_{uv} = s^{\alpha}_{uv}$ we find $t^{\beta}_{uv} = \sum_{\alpha=1}^{g} C_{\alpha\beta} s^{\alpha}_{uv}$ where $\| C_{\alpha\beta} \|$ is the reciprocal of $\| c_{\alpha\beta} \|$. To find the system of quantities of which

$$\phi_\alpha = | B^{\alpha}_{uv} |^{k\alpha/2} \cdot | B^{\alpha}_{uv} - i s^{\alpha}_{uv} |^{-(k\alpha/2)} \tag{6}$$

is the characteristic function, we express the t^{β}_{uv} in the right side of (5) in terms of the s^{β}_{uv}. Accordingly, we find that ϕ_α is the characteristic function of the quantities

$$b^{\alpha}_{uv} = \sum_{\beta} C_{\alpha\beta} \eta^{\beta}_{uv}. \tag{7}$$

Since $\phi = \prod_{\alpha=1}^{g} \phi_\alpha$ the systems $\{b^{\alpha}_{uv}\}$ $(\alpha = 1, 2, \cdots, g)$ are independently distributed. If $k_\alpha > m + 1$, the distribution of the quantities $\{b^{\alpha}_{uv}\}$ can be shown [2] from ϕ_α to be

$$V_m(B^{\alpha}_{uv}, b^{\alpha}_{uv}, k_\alpha) = \frac{| B^{\alpha}_{uv} |^{k\alpha/2} | b^{\alpha}_{uv} |^{(k\alpha-m-1)/2} \exp\left(-\sum_{u,v} B^{\alpha}_{uv} b^{\alpha}_{uv}\right)}{\pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma^{-[(k\alpha+1-i)/2]}} \prod_{u,v} db^{\alpha}_{uv} \tag{8}$$

The distributions given by (8) for all α's for which $k_\alpha > m$ are the basic

[2] J. Wishart and M. S. Bartlett, *Proceedings of the Cambridge Philosophical Society*, vol. 29 (1933), pp. 260-270.

distributions from which distributions of systems of variances, covariances and intraclass covariances can be found by transforming the $b^{\alpha}{}_{uv}$. The foregoing method can be conveniently summarized in the following

THEOREM. *If the characteristic function* $|A_{pq}|^{1/2}\,|A_{pq}-it_{pq}|^{-(1/2)}$ $(p,q=1,2,\cdots,n;\ \|A_{pq}\|$ *positive definite) of a system of second order products* $y_p y_q=\xi_{pq}$ *factors into the form*

$$\prod_{\alpha=1}^{g}\Big|\sum_{\beta}c_{\alpha\beta}A^{\beta}{}_{uv}\Big|^{k\alpha/2}\cdot\Big|\sum_{\beta}c_{\alpha\beta}(A^{\beta}{}_{uv}-it^{\beta}{}_{uv})\Big|^{-(k\alpha/2)},$$

where $u,v=1,2,\cdots m$; $|c_{\alpha\beta}|\neq 0$, *and* $\|C_{\alpha\beta}\|=\|c_{\alpha\beta}\|^{-1}$, *and the quantities* $A^{\beta}{}_{uv}$, $t^{\beta}{}_{uv}$ *and* $\eta^{\beta}{}_{uv}$ *are defined as in* (4) *and* (5), *then the systems* $\{b^{\alpha}{}_{uv}=\sum_{\beta}C_{\alpha\beta}\eta^{\beta}{}_{uv}\}$ *are independent of each other in the probability sense and if* $k_\alpha>m+1$, *the distribution of the system* $\{b^{\alpha}{}_{uv}\}$ *is given by* $V_m\,(\sum_{\beta}c_{\alpha\beta}A^{\beta}{}_{uv},\ b^{\alpha}{}_{uv},\ k_\alpha)$.

Now, as an illustration of the method, let us consider an interrelated system of families $F_1,F_2,\cdots F_m$ of k variates each, where $x_{u1},x_{u2},\cdots x_{uk}$ are the variates in F_u. Suppose the system is normally distributed such that m_u and $\sigma_u{}^2$ are the mean and variance of each variate in F_u, ρ_{uv} the correlation ($\sigma_u\sigma_v\rho_{uv}$ the covariance) between x_{ui} and x_{vi} $(i=1,2,\cdots k)$ and ρ'_{uv} the intraclass correlation ($\sigma_u\sigma_v\rho'_{uv}$ the intraclass covariance) between x_{ui} and x_{vj} $(i,j=1,2,\cdots k;\ i\neq j)$. Here, of course, $\rho_{uv}=\rho_{vu}$, $\rho'_{uv}=\rho'_{vu}$, $\rho_{uu}=1$. The elements in the symmetric and positive definite matrix Δ of correlations can be arranged so that Δ will consist of m^2 square blocks of k^2 elements each such that if θ_{uvij} is the general element in the matrix then

$$\theta_{uvij}\begin{cases}=1, & i=j,\ u=v,\\ =\rho_{uv}, & i=j,\\ =\rho'_{uv}, & i\neq j.\end{cases}\tag{9}$$

If we let the reciprocal of the matrix $\|\sigma_u\sigma_v\theta_{uvij}\|$ of variances and covariances be $\|A_{uvij}\|$ whose determinant is $|A|$, then the distribution of the x_{ui} is

$$U_{mk}(\tfrac{1}{2}A_{uvij},x_{ui}-m_u)=\frac{|A|^{1/2}}{(2\pi)^{mk/2}}\exp\Big(-\tfrac{1}{2}\Sigma A_{uvij}(x_{ui}-m_u)\,(x_{vj}-m_v)\Big)\prod_{u,i}dx_{ui}\tag{10}$$

where the sum and product are to be taken for all subscripts. It can be readily verified that A_{uvij} is of the form

$$A_{uvij} \begin{cases} = A_{uv}, & i = j \\ = A'_{uv}, & i \neq j \end{cases} \tag{11}$$

where the A_{uv} and A'_{uv} are functions of the σ's and ρ's which need not be given for our purposes. Now suppose observations have been made on N sets of families distributed according to (9). Denoting the values of the x's for the τ-th set of families by $x_{ui\tau}$ $(u = 1, 2, \cdots m;\ i = 1, 2, \cdots k)$ the probability of obtaining such a set of observed values of the x's is

$$P = \prod_{\tau=1}^{N} U_{mk}(A_{uvij}, x_{ui\tau} - m_u). \tag{12}$$

Now let us consider the product sums

$$\left.\begin{aligned} a_{uvij} &= \sum_{\tau} (x_{ui\tau} - \bar{x}_{ui})(x_{vj\tau} - \bar{x}_{vj}) \\ b_{uvij} &= N(\bar{x}_{ui} - m_u)(\bar{x}_{vj} - m_v) \end{aligned}\right\} (u, v = 1, 2 \cdots m;\ i, j = 1, 2 \cdots k) \tag{13}$$

where $\bar{x}_{ui} = (1/N) \sum_{\tau} x_{ui\tau}$. It is known from the sampling properties of normally distributed variables that the distribution of the system of means is independent of that of the second order product moments of deviations from means. The characteristic function of the a's and b's is given by the expression

$$\begin{aligned} \phi &= E[\exp(\Sigma\alpha_{uvij}a_{uvij} + \Sigma\beta_{uvij}b_{uvij})] \\ &= |\tfrac{1}{2}A|^{N/2} \, |\tfrac{1}{2}A_{uvij} - \alpha_{uvij}|^{-[(N-1)/2]} \, |\tfrac{1}{2}A_{uvij} - \beta_{uvij}|^{-(1/2)}. \end{aligned} \tag{14}$$

To adhere to the usual definition of a characteristic function the α's and β's will be purely imaginary numbers; all other numbers will be real. We now let

$$\begin{aligned} \alpha_{uvij} &\begin{cases} = \alpha_{uv}, & i = j \\ = \alpha'_{uv}, & i \neq j \end{cases} \\ \beta_{uvij} &\begin{cases} = \beta_{uv}, & i = j \\ = \beta'_{uv}, & i \neq j. \end{cases} \end{aligned} \tag{15}$$

Making use of (11) we find that ϕ becomes the characteristic function of the quantities in the square brackets in the following expression

$$\begin{aligned} \phi' = E\{\exp(&\sum_{u,v} \alpha_{uv}[\sum_{i} a_{uvii}] + \sum_{u,v} \alpha'_{uv}[\sum_{i \neq j} a_{uvij}] \\ &+ \sum_{u,v} \beta_{uv}[\sum_{i} b_{uvii}] + \sum_{u,v} \beta'_{uv}[\sum_{i \neq j} b_{uvij}])\}. \end{aligned} \tag{16}$$

Using equations (11) and (15) in (14) we find that ϕ' factors into the form

$$\begin{aligned} &|\tfrac{1}{2}A|^{N/2} \, |\tfrac{1}{2}C_{uv} - \gamma_{uv}|^{-[(N-1)(k-1)/2]} \cdot |\tfrac{1}{2}C_{uv} - \gamma'_{uv}|^{-[(k-1)/2]} \\ &\qquad \cdot |\tfrac{1}{2}D_{uv} - \delta_{uv}|^{-[(N-1)/2]} \cdot |\tfrac{1}{2}D_{uv} - \delta'_{uv}|^{-(1/2)}, \end{aligned} \tag{17}$$

where

$$(18)\quad \begin{aligned} \gamma_{uv} &= \alpha_{uv} - \alpha'_{uv},\ \delta_{uv} = \alpha_{uv} + (k-1)\alpha'_{uv},\ C_{uv} = A_{uv} - A'_{uv} \\ \gamma'_{uv} &= \beta_{uv} - \beta'_{uv},\ \delta'_{uv} = \beta_{uv} + (k-1)\beta'_{uv},\ D_{uv} = A_{uv} + (k-1)A'_{uv}. \end{aligned}$$

In terms of the σ_u, ρ_{uv} and ρ'_{uv} the matrices $\| C_{uv} \|$ and $\| D_{uv} \|$ are the reciprocals of $\| \sigma_u\sigma_v(\rho_{uv} - \rho'_{uv}) \|$ and $\| \sigma_u\sigma_v(\rho_{uv} + (k-1)\rho'_{uv}) \|$ respectively. Let the α's and β's in (16) be transformed to the γ's and δ's by the equations (18), afterwards setting $\gamma'_{uv} = \gamma_{uv}$. ϕ' becomes

$$(19)\quad \phi'' = E\{\exp(\sum_{u,v}[\gamma_{uv}c_{uv} + \delta_{uv}d_{uv} + \delta'_{uv}e_{uv}])\}$$

where

$$(20)\quad \begin{aligned} c_{uv} &= \frac{k-1}{k}\sum_i (a_{uvii} + b_{uvii}) - \frac{1}{k}\sum_{i\neq j}(a_{uvij} + b_{uvij}) \\ d_{uv} &= \frac{1}{k}\sum_{i,j} a_{uvij} \\ e_{uv} &= \frac{1}{k}\sum_{i,j} b_{uvij}. \end{aligned}$$

Thus, the value of ϕ'' in (19) is given by (17) with the γ'_{uv} replaced by γ_{uv}.

Since the characteristic function of the systems $\{c_{uv}\}$ and $\{d_{uv}\}$ is the product of the characteristic functions for the two systems, each having the form of ϕ_a as given by (6), the distributions of $\{c_{uv}\}$ and $\{d_{uv}\}$ are therefore $V_m(\frac{1}{2}C_{uv}, c_{uv}, N(k-1))$ and $V_m(\frac{1}{2}D_{uv}, d_{uv}, N-1)$ respectively. If we let

$$(21)\quad \bar{x}_u = \frac{1}{k}\sum_{i=1}^{k} \bar{x}_{ui},$$

then $e_{uv} = Nk(\bar{x}_u - m_u)(\bar{x}_v - m_v)$ and the quantities $\sqrt{Nk}(\bar{x}_u - m_u)$ are distributed according to the normal law $U_m(\frac{1}{2}D_{uv}, \sqrt{Nk}(\bar{x}_u - m_u))$ whose functional form is given by (1), and independently of the c_{uv} and d_{uv}.

Let the values of the a's and b's given by (13) be substituted in (20), and let

$$(22)\quad \begin{aligned} s_{uv} &= \sum_{a,i}(x_{uia} - \bar{x}_u)(x_{via} - \bar{x}_v) \\ s'_{uv} &= \sum_{a,i\neq j}(x_{uia} - \bar{x}_u)(x_{vja} - \bar{x}_v). \end{aligned}$$

Then

$$(23)\quad \begin{aligned} c_{uv} &= \left(1 - \frac{1}{k}\right)s_{uv} - \frac{1}{k}s'_{uv} \\ d_{uv} &= \frac{1}{k}(s_{uv} + s'_{uv}). \end{aligned}$$

The sample values of the variance σ_u^2, covariance $\sigma_u\sigma_v\rho_{uv}$ and intraclass covariance $\sigma_u\sigma_v\rho'_{uv}$ are given by

$$(24) \qquad s_u^2 = \frac{s_{uu}}{Nk}, \qquad \bar{s}_{uv} = \frac{s_{uv}}{Nk}, \qquad \bar{s}'_{uv} = \frac{s'_{uv}}{Nk(k-1)}.$$

The sample values of ρ_{uv} and ρ'_{uv} are respectively

$$(25) \qquad r_{uv} = \frac{s_{uv}}{\sqrt{s_{uu}s_{vv}}}, \qquad r'_{uv} = \frac{s'_{uv}}{(k-1)\sqrt{s_{uu}s_{vv}}}.$$

Incidentally, it should be remarked that $\bar{x}_u$, s_u^2, r_{uv}, and r'_{uv} are the optimum estimates of the parameters m_u, σ_u^2, ρ_{uv}, and ρ'_{uv}; that is, the values of the parameters which maximize P in (12) for a given set of x's.

The distribution of the $\bar{s}_{uv}$ and $\bar{s}'_{uv}$ will be given by the product $V_m(\frac{1}{2}C_{uv}, c_{uv}, N(k-1)) \cdot V_m(\frac{1}{2}D_{uv}, d_{uv}, N-1)$ after applying the transformations (23) and (24). The result is

$$(26) \quad V'_m(\tfrac{1}{2}C_{uv}, N(k-1)(\bar{s}_{uv} - \bar{s}'_{uv}), N(k-1)) \cdot V'_m(\tfrac{1}{2}D_{uv}, N(\bar{s}_{uv} + \overline{k-1}\bar{s}'_{uv}), N-1)[N^2k(k-1)]^{m(m+1)/2} \prod d\bar{s}_{uv}d\bar{s}'_{uv}$$

where $V'_m(\ )$ denotes $V_m(\ )$ with differentials omitted. Similarly, the distribution of the s_u^2, r_{uv} and r'_{uv} can be found by applying the transformations (24) and (25) to (26).

The problem of testing from the sample, that is, the information contained in the $\bar{s}_{uv}$ and $\bar{s}'_{uv}$, that the intraclass correlations ρ'_{uv} are all zero can be solved by considering the Neyman-Pearson [3] criterion for this hypothesis. The criterion is

$$(27) \qquad \lambda_{(\rho'_{uv}=0)} = \frac{|\bar{s}_{uv} - \bar{s}'_{uv}|^{N(k-1)} \cdot |\bar{s}_{uv} + (k-1)\bar{s}'_{uv}|^{N}}{|\bar{s}_{uv}|^{Nk}}$$

and is the ratio of the maximum of P in (12) for variations of the parameters m_u, σ_u^2 and ρ_{uv} (with the $\rho'_{uv} = 0$) to the maximum of P for variations of all parameters. (27) is clearly a function of generalized variances. The sampling moments and the distribution of λ, under the assumption of zero intraclass correlations, can be found by a method discussed by the author elsewhere.[4] The criterion actually used in the one variable problem ($u = 1$) to test the hypothesis that the $\rho' = 0$ is

$$(28) \qquad \theta = \frac{\bar{s}_{11} + \overline{k-1}\bar{s}'_{11}}{\bar{s}_{11} - \bar{s}'_{11}}.$$

The analogous criterion for m variables is

---

[3] J. Neyman and E. S. Pearson, *Philosophical Transactions of the Royal Society of London*, Ser. A, vol. 231 (1932), p. 295.

[4] *Annals of Mathematics*, vol. 35 (1934), p. 323.

$$(29) \qquad \frac{|\bar{s}_{uv} - \bar{s}'_{uv}|}{|\bar{s}_{uv}|}$$

which is a generalization of $\frac{k}{\theta + k - 1}$. The moments and distribution of the expression in (29) can also be found by the methods just cited.

The methods used in this paper can be applied to problems of finding the general sampling distributions of systems of variances and covariances associated with Latin Squares, "equalized" random blocks, and various other lay-outs used in experimental agriculture.

PRINCETON UNIVERSITY.

TOTALLY DISCONNECTED LOCALLY COMPACT RINGS.[1]

By N. JACOBSON.

Introduction.

1. Miss Taussky and I[2] have recently considered the theory of locally compact separable (l. c. s.) rings and shown that it could be reduced for the most part to the consideration of two essentially different topological types, the connected and the totally disconnected. Furthermore the connected l. c. s. rings satisfying mild algebraic conditions were shown to be hypercomplex systems with finite bases over the field of real numbers and hence their structure could be given by means of the classical results on hypercomplex numbers.

In this paper I consider the totally disconnected (t. d.) l. c. s. rings restricting myself in the main to simple rings and particularly to fields. The work is divided into two parts which are quite independent of each other. Part I is concerned with the nature of the additive abelian group of a locally-compact separable totally-disconnected (l. c. s. t. d.) simple ring $\mathfrak{S}$. $\mathfrak{S}$ may be non-commutative and non-associative. It is necessary to distinguish two cases: (1) the characteristic of $\mathfrak{S}$, $\chi(\mathfrak{S}) = p$ a prime and (2) $\chi(\mathfrak{S}) = 0$. In (1) the additive group of $\mathfrak{S}$ is a direct sum of cyclic groups of order p. (The precise meaning of this type of direct sum usually involving an infinite number of summands is given below (§ 5).) In (2) $\mathfrak{S}$ is additively a finite dimensional vector space over the field of p-adic numbers $\mathbf{P}_p$. It follows that $\mathfrak{S}$ is a hypercomplex system with a finite basis relative to $\mathbf{P}_p$.

Now if $\mathfrak{S}$ is associative and $\chi(\mathfrak{S}) = 0$, the above result together with the known theory of hypercomplex systems over a p-adic field gives a complete solution of the problem of determining the algebraic structure of $\mathfrak{S}$. For, by Wedderburn's theorem $\mathfrak{S}$ is a system consisting of all matrices of a fixed finite degree with coefficients in a p-adic division algebra $\mathfrak{F}$. By the results of Hasse[3] $\mathfrak{F}$ is cyclic over its centrum $\mathfrak{C}$ and the precise nature of the latter can be described.

On the other hand there is no such well-developed theory which will apply directly to solve the structure problem for the case $\chi(\mathfrak{S}) = p$. It is necessary

[1] Presented to the Society, April 19, 1935.

[2] "Locally compact rings," *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 106-108.

[3] H. Hasse, "Über $\mathfrak{P}$-adische Schiefkörper etc.," *Mathematische Annalen*, vol. 104 (1931), referred to below as H.

to develop new methods for the solution of this problem. This has been done in part II under the further assumption that $\mathfrak{S} = \mathfrak{F}$ is a field (not necessarily commutative). Moreover the methods apply at the same time to fields of characteristic 0 and thus afford a new treatment of p-adic division algebras. The main idea underlying this treatment is that the orders (Ordnungen) of a p-adic division algebra are compact and open (c. o.) subrings and conversely. Since the existence of c. o. subrings of $\mathfrak{F}$ can be shown directly the procedure is as follows: We first consider the ideal theory of any c. o. subring $\mathfrak{R}$ of $\mathfrak{F}$ and then prove the existence of a unique maximal c. o. subring $\mathfrak{R}_\omega$. By means of $\mathfrak{R}_\omega$ a valuation (Bewertung) of $\mathfrak{F}$ is defined. We then use the methods of Hasse to show that $\mathfrak{F}$ is a cyclic algebra over its centrum $\mathfrak{C}$. The structure of the valued commutative field $\mathfrak{C}$ can be determined as, indeed, it has been by v. Dantzig [4] and by Hasse and Schmidt.[5]

2. For a comprehensive account of the foundations of topological algebra the reader is referred to the paper by v. Dantzig (D. l. c. in footnote 4). We recall a few of the important definitions at this point.

By a *space* $\mathfrak{S}$ we shall always mean a Hausdorff space.[6] $\mathfrak{S}$ is said to be *compact* if every infinite sequence of points in it has a limit point. $\mathfrak{S}$ is *locally compact* (l. c.) if every point has a neighborhood whose closure is compact. $\mathfrak{S}$ is *separable* (s.) if it contains a denumerable set of neighborhoods which generate all the open sets by logical addition. $\mathfrak{S}$ is *connected* if it is impossible to decompose it into a logical sum $\mathfrak{S}_1 \cup \mathfrak{S}_2$ of the open and non-intersecting sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$. $\mathfrak{S}$ is *totally disconnected* (t. d.) if for every pair of distinct points a_1, a_2 there is a decomposition of the space into a sum of non-intersecting open sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$ such that $\mathfrak{S}_i \supset a_i$. If $\mathfrak{S}$ is l. c. and t. d. then it is *zero-dimensional*, i. e., every point has arbitrarily small open and closed neighborhoods.[7]

A *locally compact separable ring* $\mathfrak{S}$ is a l. c. s. space in which there is defined two binary operations $+$ and $\cdot$ such that

1. $\mathfrak{S}$ is an abelian group under $+$.

---

[4] D. van Dantzig in *Studien over topologische Algebra*, Dissertation, Amsterdam, H. J. Paris, 1931, determines the structure of all l. c. commutative and associative fields. The foundations of topological algebra may be found in this paper (in Dutch) or in van Dantzig, "Zur topologische Algebra," *Mathematische Annalen*, vol. 107 (1933), referred to as D.

[5] H. Hasse and F. K. Schmidt, "Die Struktur diskret bewerteter Körper," *Journal f. d. reine u. angew. Math.*, vol. 170 (1933), pp. 4-63.

[6] F. Hausdorff, *Grundzüge der Mengenlehre*, 1914, p. 213.

[7] K. Menger, *Dimensionstheorie*, Teubner, 1928, p. 207.

2. $a(b+c) = ab + ac$, $(b+c)a = ba + ca$.

3. $+$, $-$, $\cdot$, and $\div$ (when it exists) are continuous operations: If $a_\mu \to a$ (converges to a) and $b_\mu \to b$, then $a_\mu \pm b_\mu \to a \pm b$ and $a_\mu b_\mu \to ab$. If a_μ^{-1} and a^{-1} exist, then $a_\mu^{-1} \to a^{-1}$.

I. The additive group of a simple ring.

3. A l. c. s. t. d. ring $\mathfrak{S}$ is a l. c. s. t. d. abelian group under addition. It is therefore natural to begin our investigation with an account of the theory of l. c. s. t. d. abelian groups (§ 3 and § 4). It may be noted in the sequel that for our present purpose (the classification of simple rings) only a portion of the group theoretic results are needed. However, we may justify the more complete discussion on the ground that it shows up the precise extent to which the ring restrictions effect the structure of its additive group, and it should be useful for future generalizations. The main results on l. c. s. t. d. abelian groups are due to v. Dantzig (dissertation) and independently to Alexander and Cohen.[8]

In the remainder of this section $\mathfrak{S}$ will denote a l. c. s. t. d. abelian group with addition as the group operation.

THEOREM 3.1. *$\mathfrak{S}$ is complete.*

The completeness is meant in the following sense: if $a_1, a_2, \cdots$ is a Cauchy sequence, i. e. has the property that for any given neighborhood U of 0 there exists a positive integer $N(U)$ such that $a_\mu - a_\nu \in U$ if μ and $\nu > N$, then $a_1, a_2, \cdots$ converges. To prove this property suppose U is compact, i. e., has a compact closure. Since $a_{N+1} - a_N, a_{N+2} - a_N, \cdots$ all belong to U, they have a limit b. It follows easily that $a_\mu \to a_N + b$.

THEOREM 3.2. *If U is an open and closed compact neighborhood of 0, then there exists an open and closed compact subgroup $\mathfrak{G} \subset U$.*[9]

Denote the set of elements $\{-u\}$ where $u \in U$ by $-U$ and $V = U \cap (-U)$ (logical intersection). Since $-U$ and U are open and closed, so is their intersection V. We denote the complement of V in $\mathfrak{S}$ by $\mathfrak{S} \mid V$.

---

[8] J. W. Alexander and L. W. Cohen in "A classification of the homology groups of compact spaces," *Annals of Mathematics*, vol. 33 (1932), pp. 538-566, deal with groups with generators, which may be shown by virtue of Theorem 3.2 to include the c. s. t. d. groups.

[9] This theorem is due to v. Dantzig cf. his dissertation, p. 18, and also E. R. v. Kampen, "Locally compact abelian groups," *Proceedings of the National Academy of Sciences*, vol. 20 (1934).

If $U_1 \supset U_2 \supset U_3 \supset \cdots$ is a decreasing sequence of neighborhoods of 0 whose intersection is 0 ($U_\mu \to 0$) then for μ sufficiently large $(U_\mu + V) \cap (\mathfrak{S} \mid V)$ is vacuous. Otherwise we have for every μ a $u_\mu \in U_\mu$ and $v_\mu \in V$ such that $u_\mu + v_\mu = h_\mu \in \mathfrak{S} \mid V$. Suppose $v_{\mu_k} \to v \in V$. Since $u_{\mu_k} \to 0$, $h_{\mu_k} \to v$ which is impossible since $\mathfrak{S} \mid V$ is closed.

Let $\mathfrak{G}$ denote the set of elements g of $\mathfrak{S}$ such that $g + V \subset V$. $\mathfrak{G}$ is a group $\subset V$ and by the above remark $\mathfrak{G} \supset U_\mu$ for μ sufficiently large. Hence $\mathfrak{G}$ is open. Being a group $\mathfrak{G}$ is also closed.[10]

By Theorem 3.2 there exists a sequence of compact open and closed subgroups $\mathfrak{G}_\mu \to 0$. The cosets of $\mathfrak{G}_\mu$ constitute a fundamental set of neighborhoods for $\mathfrak{S}$. In the rest of I we will therefore mean by a neighborhood of 0 a compact open and closed group neighborhood.

COROLLARY. $\sum_{\mu=1}^{\infty} a_\mu$ *exists if and only if* $a_\mu \to 0$.

Set $\sum_{\mu=1}^{n} a_\mu = s_n$. By the completeness of $\mathfrak{S}$, $\sum_{\mu=1}^{\infty} a_\mu$ exists if and only if $\{s_n\}$ is a Cauchy sequence. Let U be a (compact group) neighborhood of 0. If $\mu > N(U)$, $a_\mu \in U$ and hence also for $n, m > N$

$$s_n - s_m = \begin{cases} a_{m+1} + a_{m+2} + \cdots + a_n & \text{if } n > m \\ 0 & \text{if } n = m \\ a_{n+1} + a_{n+2} + \cdots + a_m & \text{if } n < m \end{cases}$$

is contained in U. Thus $\{s_n\}$ is a Cauchy sequence and Σa_μ exists. The converse is obvious.

4. A group $\mathfrak{S}$ is called a *p-group* (p a prime) if for every $a \in \mathfrak{S}$, $p^\mu a \to 0$. Let $\mathfrak{G}$ be a compact t. d. group and

$$\mathfrak{G}^{(p)} = \underset{\mu}{I}(p^\mu \mathfrak{G}) \equiv \mathfrak{G} \cap p\mathfrak{G} \cap p^2\mathfrak{G} \cap \cdots .$$

LEMMA 4.1. *A necessary and sufficient condition that an element* $h \in \mathfrak{G}^{(p)}$ *is that there exist a sequence of integers* $\mu_i \to \infty$ *such that* $p^{\mu_i} h \to h$.

(1) Suppose $p^{\mu_i} h \to h$. Since $p^{\mu_i} h \in p^\mu \mathfrak{G}$ for $\mu_i \geqq \mu$ and $p^\mu \mathfrak{G}$ is closed, we have $p^\mu \mathfrak{G} \supset h$. Hence $h \in \underset{\mu}{I} p^\mu \mathfrak{G} = \mathfrak{G}^{(p)}$.

(2) Let $h \in \mathfrak{G}^{(p)}$ so that $h = p^\nu g_\nu$ $(\nu = 1, 2, \cdots)$. If $g_{\nu_k} \to g$ then $h = \lim_{k \to \infty} p^{\nu_k} g$. For if k is sufficiently large $(g_{\nu_k} - g) \in U$ where U is an arbitrary (group) neighborhood of 0. Hence

[10] D. p. 609. That $\mathfrak{G}$ is closed may also be seen directly.

$$p^{\nu_k}(g_{\nu_k} - g) = (h - p^{\nu_k}g) \in U$$

or $\lim_{k\to\infty} p^{\nu_k}g = h$. Set $\mu_k = \nu_{k+1} - \nu_k$. By dropping enough terms of the sequence $\{\nu_k\}$ we may suppose that $\mu_k \to \infty$. We assert that $p^{\mu_k}h \to h$. If k is large enough, $(h - p^{\nu_k}g) \in U$ and hence

$$p^{\mu_k}(h - p^{\nu_k}g) = (p^{\mu_k}h - p^{\nu_{k+1}}g) \in U$$
$$(h - p^{\mu_k}h) = (h - p^{\nu_{k+1}}g) - (p^{\mu_k}h - p^{\nu_{k+1}}g) \in U$$

i. e. $p^{\mu_k}h \to h$.

LEMMA 4. 2. *If* $p^{\mu_k}g \to 0$ *then* $p^{\mu}g \to 0$ $(\mu = 1, 2, \cdots)$.

If k is large enough so that $p^{\mu_k}g \in U$ then $p^{\mu}g \in U$ also for $\mu \geqq \mu_k$.

According to this lemma we have that the elements g for which the sequence $\{p^{\mu}g\}$ has 0 for limit point form a closed subgroup $\mathfrak{G}_p$.

LEMMA 4. 3. $\mathfrak{G} = \mathfrak{G}^{(p)} \oplus \mathfrak{G}_p$.

Let g be an arbitrary element of $\mathfrak{G}$ and $p^{\nu_k}g \to h$. Since $p^{\mu_k}h \to h$ where $\mu_k = \nu_{k+1} - \nu_k$, $h \in \mathfrak{G}^{(p)}$. Suppose $p^{\mu'_k}g \to h' \in \mathfrak{G}^{(p)}$. Then $p^{\nu'_k}h' \to h$. For if U is a neighborhood of 0 and $k > K_1(U)$, $(h' - p^{\mu'_k}g) \in U$ and hence also $(p^{\nu'_k}h' - p^{\nu'_{k+1}}g) \in U$. If $k > K_2(U)$ we have besides $(p^{\nu'_{k+1}}g - h) \in U$. Hence

$$(h - p^{\nu'_k}h') \in U, \quad \text{or} \quad p^{\nu'_k}h' \to h.$$

Thus $p^{\nu'_k}(g - h') \to 0$ and by Lemma 4. 2 $p^{\nu}(g - h') \to 0$.

Set $g = h' + (g - h')$. $h' \in \mathfrak{G}^{(p)}$ and $(g - h') \in \mathfrak{G}_p$. Evidently this decomposition is unique, i. e., $\mathfrak{G}^{(p)} \cap \mathfrak{G}_p = 0$ and so $\mathfrak{G} = \mathfrak{G}^{(p)} \oplus \mathfrak{G}_p$. (This shows that h' determined above is unique and in particular $p^{\mu_k}g \to h'$ not merely $p^{\mu'_k}g$).

LEMMA 4. 4. *There exist primes* p *such that* $\mathfrak{G}^{(p)} \neq \mathfrak{G}$ *and hence such that* $\mathfrak{G}_p \neq 0$.

If $\mathfrak{H}_0$ is an open subgroup of $\mathfrak{G}$ then $\mathfrak{G} - \mathfrak{H}_0$ is compact and discrete and hence it is finite. Let p be a divisor of its order. $\mathfrak{G} - \mathfrak{H}_0$ has a subgroup $\mathfrak{H} - \mathfrak{H}_0$ of index p. Hence $\mathfrak{G} - \mathfrak{H} \cong (\mathfrak{G} - \mathfrak{H}_0) - (\mathfrak{H} - \mathfrak{H}_0)$ has order p and $p\mathfrak{G} \subset \mathfrak{H} < \mathfrak{G}$ ($<$ means properly contained in). It follows from Lemma 4. 3 that $\mathfrak{G}_p \neq 0$.

THEOREM 4. 1. $\mathfrak{G}$ *is a direct sum of p-groups.*[11]

[11] Alexander and Cohen, *loc. cit.*,[8] p. 557.

Let p_1 be a prime for which $\mathfrak{G}_{p_1} \neq 0$. We have $\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}$. Since $\mathfrak{G}^{(p_1)}$ is closed, it is compact and may be treated as $\mathfrak{G}$, i. e. $\mathfrak{G}^{(p_1)} = \mathfrak{G}^{(p_1)}_{p_2} \oplus \mathfrak{G}^{(p_1 p_2)}$. Continuing this process we obtain

$$\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}_{p_2} \oplus \cdots \oplus \mathfrak{G}^{(p_1 \cdots p_{n-1})}_{p_n} \oplus \mathfrak{G}^{(p_1 \cdots p_n)}$$

where $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}} \neq 0$ for $i < n$ if $\mathfrak{G}^{(p_1 \cdots p_n)} \neq 0$. We must show that this process is uniformly convergent in the sense that $\mathfrak{G}^{(p_1 \cdots p_n)} \subset U$ where U is any neighborhood of 0 and $n > N(U)$. Choose N so that $\{p_1, p_2, \cdots, p_n\}$ for $n > N$ includes all the distinct prime factors of the order m of $\mathfrak{G} - U$. Let $h \in \mathfrak{G}^{(p_1 \cdots p_n)}$. Then there exist sequences of integers $\{\mu_k\}, \{\nu_k\}, \cdots, \{\rho_k\}$ such that $p_1^{\mu_k} h \to h, p_2^{\nu_k} h \to h, \cdots, p_n^{\rho_k} h \to h$. It follows that $p_1^{\mu_k} p_2^{\nu_k} \cdots p_n^{\rho_k} h \to h$. If $\mu_k, \nu_k, \cdots, \rho_k$ are sufficiently large $p_1^{\mu_k} p_2^{\nu_k} \cdots p_n^{\rho_k} h = p_1^{\mu'_k} p_2^{\nu'_k} \cdots p_n^{\rho'_k} m h \in m\mathfrak{G}$ and hence $h \in m\mathfrak{G} \subset U$ or $\mathfrak{G}^{(p_1 \cdots p_n)} \subset U$. Hence we may write

$$\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}_{p_2} \oplus \mathfrak{G}^{(p_1 p_2)}_{p_3} \oplus \cdots.$$

The components $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}}$ are *characteristic subgroups* of $\mathfrak{G}$, i. e. they are carried into themselves by every automorphisms of $\mathfrak{G}$.

5. Let $\mathfrak{S}$ be a l. c. s. t. d. simple [12] ring (not necessarily associative or commutative). Since $\mathfrak{S}$ is homogeneous it is either discrete or dense in itself. The first case is, of course, uninteresting for the present considerations and hence we restrict ourselves to the second.

THEOREM 5.1. *$\mathfrak{S}$ is a p-group.*

Let $\mathfrak{S}_p$ denote the subset of elements a of $\mathfrak{S}$ such that $p^\nu a \to 0$. $\mathfrak{S}_p$ is an ideal [13] and hence either $\mathfrak{S}_p = \mathfrak{S}$ or $\mathfrak{S}_p = 0$. $\mathfrak{S}$ contains a compact and open subgroup $\mathfrak{G}$. Since $\mathfrak{S}$ is not discrete $\mathfrak{G} \neq 0$. By Lemma 4.4 $\mathfrak{G}$ contains elements $g \neq 0$ such that $p^\nu g \to 0$ for some p. Thus $\mathfrak{S}_p \neq 0$ and so $\mathfrak{S}_p = \mathfrak{S}$.

p is unique. For, suppose that $p_1^\nu a \to 0$ and $p_2^\nu a \to 0$ where $(p_1, p_2) = 1$. Integers α_ν and β_ν can be determined so that $\alpha_\nu p_1^\nu + \beta_\nu p_2^\nu = 1$ $(\nu = 1, 2, 3, \cdots)$. If U is an arbitrary neighborhood of 0, there exists an $N(U)$ such that $p_1^\nu a$ and $p_2^\nu a \in U$ for all $\nu \geqq N$. Then $a = \alpha_\nu p_1^\nu a + \beta_\nu p_2^\nu a \in U$. Since U is arbitrary a must be 0.

We now distinguish two cases:

(1) *Characteristic p.* $p\mathfrak{S} = 0$.

Let $\mathfrak{G}_1 > \mathfrak{G}_2 > \mathfrak{G}_3 > \cdots$ be a sequence of compact and open subgroups of $\mathfrak{S}$ whose intersection is 0. $\mathfrak{G}_\mu - \mathfrak{G}_{\mu+1}$ has order a power of p and hence

[12] A ring $\mathfrak{S}$ is said to be *simple* if (0) and (1) $\equiv \mathfrak{S}$ are its only (two-sided) ideals.

[13] The term *ideal* will refer exclusively to two-sided ideals.

by inserting a finite number of compact and open subgroups between $\mathfrak{G}_\mu$ and $\mathfrak{G}_{\mu+1}$ and changing the notation, we may suppose that $\mathfrak{G}_\mu - \mathfrak{G}_{\mu+1}$ has order p. If x_μ is an element of $\mathfrak{G}_\mu | \mathfrak{G}_{\mu+1}$ then $\mathfrak{G}_\mu = (x^\mu) \bigoplus \mathfrak{G}_{\mu+1}$ where (x_μ) denotes the group of order p generated by x_μ. Hence $\mathfrak{G}_1 = (x_1) \bigoplus (x_2) \bigoplus \cdots \bigoplus x_{n-1} \bigoplus \mathfrak{G}_n$ and since $\mathfrak{G}_n \to 0$

$$\mathfrak{G}_1 = (x_1) \bigoplus (x_2) \bigoplus (x_3) \bigoplus \cdots .$$

Suppose $y_1, y_2, y_3, \cdots$ is a denumerable dense set in $\mathfrak{S}$. Then $\mathfrak{S} = (\mathfrak{G}_1 + y_1) \cup (\mathfrak{G}_1 + y_2) \cup \cdots$.[14] If y_{μ_1} is the first $y \nsubseteq \mathfrak{G}_1$ set $y_{\mu_1} = x_{-1}$ and form $\mathfrak{G}_{-1} = (x_{-1}) \bigoplus \mathfrak{G}_1$. Again if y_{μ_2} is the first $y \nsubseteq \mathfrak{G}_{-1}$ set $y_{\mu_2} = x_{-2}$ and form $\mathfrak{G}_{-2} = (x_{-2}) \bigoplus \mathfrak{G}_{-1}$. Continuing in this way we obtain the theorem:

THEOREM 5.2. *$\mathfrak{S}$ has a denumerable set of generators $x_1, x_2, \cdots$ and $x_{-1}, x_{-2}, \cdots$ such that every element is expressible uniquely as an infinite linear combination of the x's with coefficients mod p and in which only a finite number of the $x_{-\mu}$'s occur. Every such expression is an element of $\mathfrak{S}$.*

(2) *Characteristic* 0. $p\mathfrak{S} \neq 0$.

LEMMA 5.1. *$\mathfrak{S}$ is a hypercomplex system over the field of rational numbers* P.[15]

If m is a positive integer, then $m\mathfrak{S}$ is an ideal and hence is either $= \mathfrak{S}$ or $= 0$. But $m\mathfrak{S} = 0$ is impossible: If $m = kl$, $m\mathfrak{S} = k(l\mathfrak{S}) = 0$ and hence either $k\mathfrak{S} = 0$ or $l\mathfrak{S} = 0$. We obtain in this way that a prime q exists for which $q\mathfrak{S} = 0$. Since p is the only prime such that $p^\nu a \to 0$ for every a in $\mathfrak{S}$, $p = q$ and we have a contradiction of the assumption $p\mathfrak{S} \neq 0$.

The elements of order m form an ideal $\neq \mathfrak{S}$. Hence this ideal is $= (0)$, i. e., every element of $\mathfrak{S}$ has infinite order. From $m\mathfrak{S} = \mathfrak{S}$ follows that for every a there exists a unique a' such that $ma' = a$. We denote a' by $(1/m)a$ and define $(n/m)a = na'$ where n is any integer. If $n/m = n_1/m_1$, then $(n/m)a = (n_1/m_1)a$. Using the fact that $\mathfrak{S}$ has no elements of finite order we can easily verify the following rules:

$$(*) \qquad \begin{gathered} \alpha(a+b) = \alpha a + \alpha b \\ (\alpha+\beta)a = \alpha a + \beta a \qquad (\alpha\beta)a = \alpha(\beta a) \\ \alpha(ab) = (\alpha a)b = a(\alpha b) \end{gathered}$$

where α and β are rational numbers.

---

[14] This is the only point in the entire discussion at which the assumption of separability seems to be necessary. In the rest of the paper it may be replaced by the weaker Hausdorff first denumerability axiom.

[15] $\mathfrak{S}$ does not have a finite basis over P, however.

Let $\mathfrak{G}$ be a compact and open subgroup of $\mathfrak{S}$. Then $\mathfrak{G} > p\mathfrak{G} > p^2\mathfrak{G} > \cdots$ and $I(p^\nu\mathfrak{G}) = 0$. For otherwise by Lemma 4. 1 $\mathfrak{G}$ would contain an element $h \neq 0$ for which the sequence $\{p^\nu h\}$ has h for a limit point. Let $p^{-\mu}\mathfrak{G}$ denote the compact and open subgroup of p^μ-th parts of the elements of $\mathfrak{G}$ and $\Gamma_\mu = p^\mu\mathfrak{G} | p^{\mu+1}\mathfrak{G}$ $(\mu = 0, \pm 1, \pm 2, \cdots)$. Then $\mathfrak{G} < p^{-1}\mathfrak{G} < p^{-2}\mathfrak{G} < \cdots$. Since $p^\nu a \to 0$ for every a, there exists an integer N such that $p^N a \in \mathfrak{G}$ and hence $a \in p^{-N}\mathfrak{G}$. Thus $\mathfrak{G} \cup p^{-1}\mathfrak{G} \cup p^{-2}\mathfrak{G} \cup \cdots = \mathfrak{S}$.

If $a \in \Gamma_\mu$, evidently $p^\nu a \in \Gamma_{\mu+\nu}$ for any integral ν. If α is a rational number prime to p (numerator and denominator in the reduced form of α are prime to p), then $\alpha a \in \Gamma_\mu$. For, if α is an integer prime to p, $\alpha a \in \Gamma_{\mu+\epsilon}$ where $\epsilon \geqq 0$. But $\epsilon > 0$ is impossible since in this case $\alpha a \in p^{\mu+1}\mathfrak{G}$ contradicting the fact that the order of the coset of a in $p^\mu\mathfrak{G} - p^{\mu+1}\mathfrak{G}$ is p. The result for reciprocals of integers follows from symmetry and hence it follows for all rational α's by combining the two special cases.

If $a \in \Gamma_\mu$ we denote the real-valued function $p^{-\mu}$ of a by $\| a \|$. We have then:

$$\| a + b \| \leqq \max (\| a \|, \| b \|)$$
$$\| p^\nu a \| = p^{-\nu} \| a \|$$
$$\| \alpha a \| = \| a \| \quad \text{if} \quad (\alpha, p) = 1.$$

If we set dist $(a, b) = \| a - b \|$, we obtain a metric which gives the topology of $\mathfrak{S}$.

LEMMA 5. 2. *If $\{\alpha_\nu\}$ is a sequence of rational numbers, $\{\alpha_\nu a\}$ $(a \neq 0)$ converges if and only if $\{\alpha_\nu\}$ converges in the p-adic topology.*[16]

Because of the completeness of $\mathfrak{S}$ and of the p-adic field P_p it is necessary to show only that $\alpha_\nu a \to 0$ if and only if $\alpha_\nu \underset{p}{\to} 0$ (in the p-adic topology). This is evident from the above considerations.

LEMMA 5. 3. *$\mathfrak{S}$ is a hypercomplex system over the field of p-adic numbers.*

If α is a p-adic number, say $\alpha = \lim_p \alpha_\nu$ where the α_ν are rational numbers, then we define $\alpha a = \lim \alpha_\nu a$. This is independent of the particular sequence $\{\alpha_\nu\}$ approaching α. From the continuity of addition and multiplication we obtain the validity of equations (*) for all p-adic α, β. Finally, if $\{\alpha_\nu\}$ is a sequence of p-adic numbers converging to α and $a_\nu \to a$ then $\alpha_\nu a_\nu \to \alpha a$, i. e. αa is a continuous function of α and a.

---

[16] For a definition of the p-adic field P_p and the p-adic topology of P see v. d. Waerden, *Moderne Algebra* I, Springer, 1930, pp. 218-220.

We denote the p-adic vector space determined by the elements $a_1, a_2, \cdots, a_r$ of $\mathfrak{S}$ by $(a_1, a_2, \cdots, a_r)$.

LEMMA 5.4. *$(a_1, a_2, \cdots, a_r)$ is closed.*

We may suppose that $a_1, \cdots, a_r$ are linearly independent (with p-adic coefficients). Let $b_\nu = \beta_{\nu 1}a_1 + \cdots + \beta_{\nu r}a_r \to b$. We must show that $b \in (a_1, a_2, \cdots, a_r)$. For each ν choose $\lambda(\nu) = 1, 2, \cdots, r$ such that $\| \beta_{\nu,\lambda(\nu)}a_{\lambda(\nu)} \| \geqq \| \beta_{\nu,\mu}a_\mu \|$ for $\mu \neq \lambda$. Since $\lambda(\nu)$ has only a finite range one of its values, say $\lambda(\nu) = 1$ occurs infinitely often. By restricting ourselves to a subsequence we may suppose that $\| \beta_{\nu 1}a_1 \| \geqq \| \beta_{\nu\mu}a_\mu \|$ for all ν and μ. Then $\| (\| \beta_{\nu 1}a_1 \|)\beta_{\nu\mu}a_\mu \| \leqq 1$ and hence we may suppose that $\lim_{\nu\to\infty} \| \beta_{\nu 1}a_1 \| \beta_{\nu\mu}a_\mu$ exists and $= \gamma_\mu a_\mu$. Since $\| (\| \beta_{\nu 1}a_1 \|)\beta_{\nu 1}a_1 \| = 1$, $\gamma_1 \neq 0$. If $\lim_p \| \beta_{\nu 1}a_1 \| = \infty$, then $\| \beta_{\nu 1}a_1 \| b \to 0$ and hence $\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_r a_r = 0$ contrary to the linear independence of the a's. Thus $\lim_p \| \beta_{\nu 1}a_1 \|$ exists and since $\| \beta_{\nu 1}a_1 \| \geqq \| \beta_{\nu\mu}a_\mu \|$ we may suppose that $\beta_{\nu\mu}a_\mu \to \beta_\mu a_\mu$. Then

$$b = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_r a_r \in (a_1, a_2, \cdots, a_r).$$

THEOREM 5.3. *$\mathfrak{S}$ is a hypercomplex system with a finite basis over the field of p-adic numbers.*

Choose a_1 in Γ_0. If $\mathfrak{S} \neq (a_1)$, then there exists an $a_2 \in \Gamma_0$ such that $a_2 \nless (a_1) + p\mathfrak{G}$. For if every b in Γ_0, $\in (a_1) + p\mathfrak{G}$ then every b_ν in $\Gamma_\nu \in (a_1) + p^{\nu+1}\mathfrak{G}$. Thus

$$\begin{array}{llll} b = \alpha_1 a_1 + b_1 & b_1 \in \Gamma_{m_1} & & m_1 > 0 \\ b_1 = \alpha_2 a_1 + b_2 & b_2 \in \Gamma_{m_2} & & m_2 > m_1 \\ \cdots\cdots\cdots & \cdots\cdots & & \cdots \\ b_\nu = \alpha_\nu a_1 + b_{\nu+1} & b_{\nu+1} \in \Gamma_{m_{\nu+1}} & & m_{\nu+1} > m_\nu \end{array}$$

and hence $b = (\alpha_1 + \cdots + \alpha_\nu)a_1 + b_{\nu+1}$. Since $\lim b_\nu = 0$ and (a_1) is closed, $b \in (a_1)$ contradicting $(a_1) \neq \mathfrak{S}_1$. Now choose $a_2 \begin{cases} \in \Gamma_0 \\ \nless (a_1) + p\mathfrak{G}. \end{cases}$ The subgroup generated by a_1 and a_2 mod $p\mathfrak{G}$ then has order p^2. If $\mathfrak{S} \neq (a_1, a_2)$, we repeat the process and obtain an $a_3 \in \Gamma_0$ such that the order of the group generated by a_1, a_2, a_3 mod $p\mathfrak{G}$ is p^3 and so on. Since $\mathfrak{G} - p\mathfrak{G}$ is finite this process breaks off and we obtain a finite basis for $\mathfrak{S}$.

Before leaving this part we note that the above methods are applicable to cases other than the one we have considered. We shall not attempt to give

the maximum generality of this group-theoretic method but mention certain extensions which are readily made.

A compact t. d. ring $\mathfrak{G}$ *is a direct sum of p-rings.*

This follows at once from the remark that the $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}}$ in Theorem 4. 1 are characteristic subgroups. For, then they are ideals and the sum is direct in the sense of rings as well as of groups.

Theorem 5. 3 may be stated under the more general assumptions that $\mathfrak{S}$ is a hypercomplex system over P and is a p-group.

II. The structure of associative fields.

6. Let $\mathfrak{F}$ be a l. c. s. t. d. associative field (not necessarily commutative). For the present we may drop the distinction between the two types characteristic 0 and characteristic p. We assume, of course, that $\mathfrak{F}$ is not discrete.

There exists a $y \neq 0$ sufficiently near 0 such that $y\bar{U} < U$ where U is a compact neighborhood of 0 and $\bar{U}$ is its closure. Hence $U > y\bar{U} > y^2\bar{U} > \cdots$. Suppose $y^{\nu_k} \to z \neq 0$. Then $y^{\mu_k} \to 1$ where $\mu_k = \nu_{k+1} - \nu_k$. If u is any element of U, $y^{\mu}u \in y^{\mu_k}\bar{U}$ for $\mu \geqq \mu_k$ and hence $u = \lim_{i\to\infty} y^{\mu_i}u$ is contained in the closed set $y^{\mu_k}\bar{U}$. Thus $U \subset y^{\mu_k}\bar{U}$ which is impossible. Hence $y^{\nu} \to 0$.

We require the existence of such a y in proving

LEMMA 6. 1. *If* $a_\nu \to \infty$ *then* $a_\nu^{-1} \to 0$.[17]

Evidently $\{a_\nu^{-1}\}$ can have no limit point other than 0 and hence it is sufficient to show that 0 is really a limit point of this sequence. Let $y \neq 0$ be an element such that $y^{\nu} \to 0$ and U a compact neighborhood of 0. If infinitely many $a_\nu y^j$ for $\nu = 1, 2, \cdots$ and j fixed were contained in U, then the corresponding a_ν would be contained in the compact set $\bar{U}y^{-j}$ and hence would have a limit point. It follows that a sequence $\{b_\nu\}$ may be extracted from $\{a_\nu\}$ so that $b_\nu y^{\nu} \in \mathfrak{F}|U$. Since $y^{\mu} \to 0$ we may determine for each ν an integer k_ν such that $b_\nu y^{k_\nu} \in \mathfrak{F}|U$ but $b_\nu y^{k_\nu+1} \in U$. Evidently k_ν $(\geq \nu) \to \infty$. Because of the compactness of U we may suppose that $b_\nu y^{k_\nu+1} \to z$. Then $b_\nu y^{k_\nu} \to zy^{-1} = w \in \mathfrak{F}|U$ and hence is $\neq 0$. From $y^{-k_\nu}b_\nu^{-1} \to w^{-1}$ and $y^{k_\nu} \to 0$ follows $b_\nu^{-1} \to 0$.

Let $\mathfrak{G}$ be a compact and open subgroup of $\mathfrak{F}$ and $\mathfrak{N}(\mathfrak{G})$ the set of elements a of $\mathfrak{F}$ such that $a\mathfrak{G} \subset \mathfrak{G}$.

---

[17] $a_\nu \to \infty$ means that $\{a_\nu\}$ is absolutely divergent, i. e., has no convergent subsequence. Lemma 6. 1 is the "Perfektisierungsaxiom" for fields of van Dantzig (D, p. 612). The present proof holds for all locally compact fields satisfying the first denumerability axiom.

LEMMA 6.2. *$\mathfrak{R}(\mathfrak{G})$ is a compact and open domain of integrity (d. o. i.).*

That $\mathfrak{R}(\mathfrak{G})$ is a ring may be verified directly. Since $\mathfrak{F}$ is a field $\mathfrak{R}$ has no zero-divisors. By the continuity of multiplication we have that $\mathfrak{R}$ contains a neighborhood of 0 and hence is open. If $\{a_\nu\}$ is an infinite sequence of elements of $\mathfrak{R}$ and $g \neq 0$ belongs to $\mathfrak{G}$ then $\{a_\nu g\}$ is a sequence in $\mathfrak{G}$ and hence has a limit point h. Thus $\{a_\nu\}$ has hg^{-1} as limit point and $\mathfrak{R}$ is compact.

7. In this section we consider the ideal theory of a compact and open (c. o.) d. o. i. $\mathfrak{R}$ contained in $\mathfrak{F}$. The existence of such subrings has just been shown.

THEOREM 7.1. *$\mathfrak{R}$ satisfies the chain conditions for left-(right-) ideals.*[18]

If $\mathfrak{J}$ is a left-(right-) ideal then $\mathfrak{J} \supset \mathfrak{R}b$ where $b \in \mathfrak{J}$. If $b \neq 0$, $\mathfrak{R}b$ is open and hence $\mathfrak{J} = \sum_{b \neq 0} \mathfrak{R}b$ is open. Let $\mathfrak{J}_1 < \mathfrak{J}_2 < \mathfrak{J}_3 < \cdots$ be an increasing sequence of ideals and $\nu_1, \nu_2, \nu_3, \cdots$ the orders of the finite quotient groups $\mathfrak{R} - \mathfrak{J}_1, \mathfrak{R} - \mathfrak{J}_2, \mathfrak{R} - \mathfrak{J}_3, \cdots$. Then $\nu_1 > \nu_2 > \nu_3 > \cdots$. Hence the sequence of $\mathfrak{J}$'s is finite. Similarly, every decreasing sequence of left-(right-) ideals $\mathfrak{J}_1 > \mathfrak{J}_2 > \mathfrak{J}_3 > \cdots > \mathfrak{J}_\omega$ containing a fixed ideal $\mathfrak{J}_\omega$ is finite.

Let $\mathfrak{P}$ be the totality of elements b of $\mathfrak{R}$ for which $b\mathfrak{R} < \mathfrak{R}$. We propose to show that $\mathfrak{P}$ is a prime ideal in $\mathfrak{R}$ and that $\mathfrak{P} > \mathfrak{P}^2 > \cdots \to 0$.

LEMMA 7.1. *$\mathfrak{R}|\mathfrak{P}$ contains arbitrarily small ideals.*

If U is any neighborhood of 0, there exists a $z \neq 0$ sufficiently near 0 such that $\mathfrak{R}z\mathfrak{R} \subset U$. $\mathfrak{R}z\mathfrak{R}$ is evidently an ideal.

LEMMA 7.2. *$\mathfrak{R}|\mathfrak{P}$ is a compact group relative to multiplication.*

$\mathfrak{R}|\mathfrak{P}$ consists of those elements $a, a', \cdots$ of $\mathfrak{F}$ which satisfy $a\mathfrak{R} = \mathfrak{R}$, $a'\mathfrak{R} = \mathfrak{R}, \cdots$. It follows then that $a'a\mathfrak{R} = \mathfrak{R}$, $a^{-1}\mathfrak{R} = \mathfrak{R}$, i. e., aa', $a^{-1} \in \mathfrak{R}|\mathfrak{P}$. Hence $\mathfrak{R}|\mathfrak{P}$ is a group. Since $\mathfrak{R}|\mathfrak{P}$ is closed and contained in $\mathfrak{R}$ it is compact.

LEMMA 7.3. *An element b of $\mathfrak{R}$ belongs to $\mathfrak{P}$ if and only if $b^\nu \to 0$.*

Since b has an inverse and $\mathfrak{R} > b\mathfrak{R}$ we have $\mathfrak{R} > b\mathfrak{R} > b^2\mathfrak{R} > \cdots$. If $b^{\nu_k} \to c \neq 0$, then $b^{\mu_k} \to 1$ where $\mu_k = \nu_{k+1} - \nu_k$. But this is impossible for the reasons given in § 6. The converse is obvious.

[18] The conditions referred to are the "Teilerkettensatz" and the "eingeschrankte Vielfachenkettensatz" of E. Noether, cf. "Abstrakter Aufbaue der Idealtheorie in algebraischen Zahl-und Funktionsenkorper," *Mathematische Annalen*, vol. 96 (1926), p. 26.

If $\mathfrak{P}_1$ is the set of b's such that $\mathfrak{R}b < \mathfrak{R}$, we may show as in Lemma 7.3 that $\mathfrak{P}_1$ consists of those elements b of $\mathfrak{R}$ whose power sequence $b^\nu \to 0$. Hence $\mathfrak{P}_1 = \mathfrak{P}$.

LEMMA 7.4. *If $\mathfrak{A}$ is a finite ring, its radical $\mathfrak{N}$ consists of the properly nilpotent elements,* i. e., the elements z such that za and az are nilpotent for all a in $\mathfrak{A}$.

This lemma is well known for rings with a finite basis over a field and it may be proved in exactly the same manner for our case of finite rings.[19]

THEOREM 7.2. *$\mathfrak{P}$ is a prime ideal in $\mathfrak{R}$ and $I_\nu \mathfrak{P}^\nu = 0$.*[20]

If $b \in \mathfrak{P}$ and $a \in \mathfrak{R}$ then ab and $ba \in \mathfrak{P}$. Hence $\mathfrak{P}$ is invariant. Since $\mathfrak{P}$ contains a sufficiently small neighborhood of 0, it contains an ideal $\mathfrak{J}$ of $\mathfrak{R}$. Consider the finite ring $\mathfrak{A} = \mathfrak{R} - \mathfrak{J}$. In the isomorphism $\mathfrak{R} \sim \mathfrak{A}$ the elements of $\mathfrak{P}$ are precisely those elements of $\mathfrak{R}$ which correspond to properly nilpotent elements of $\mathfrak{A}$. Hence $\mathfrak{P}$ corresponds to the radical $\mathfrak{N}$ of $\mathfrak{A}$. It follows that $\mathfrak{P}$ is an ideal and $\mathfrak{P}^n \equiv 0 (\mathfrak{J})$ for sufficiently high n. Since $\mathfrak{J}$ is arbitrarily small, $I_\nu \mathfrak{P}^\nu = 0$. $\mathfrak{P}$ is evidently prime.

From this theorem and Wedderburn's theorem on finite fields [21] we obtain

THEOREM 7.3. *$\mathfrak{R} - \mathfrak{P}$ is a commutative field.*

8. Let $\mathfrak{R}_1$ be a c. o. d. o. i. contained in $\mathfrak{F}$ and $\mathfrak{P}_1$ its prime ideal defined as in the last section. (It is easily seen that $\mathfrak{P}_1$ is the only prime ideal in $\mathfrak{R}_1$). Consider $\mathfrak{R}_2 = \mathfrak{R}(\mathfrak{P}_1)$ the largest d. o. i. in which $\mathfrak{P}_1$ is contained as a left-ideal, i. e., $\mathfrak{R}_2$ consists of the elements a of $\mathfrak{F}$ such that $a\mathfrak{P}_1 \subset \mathfrak{P}_1$. Evidently $\mathfrak{R}_2 \supset \mathfrak{R}_1$ and is c. o. Determine $\mathfrak{R}_3 = \mathfrak{R}(\mathfrak{P}_2)$ where $\mathfrak{P}_2$ is the prime ideal of $\mathfrak{R}_2$ and so on. We thus obtain a sequence of c. o. d. o. i. $\mathfrak{R}_1 \subset \mathfrak{R}_2 \subset \mathfrak{R}_3 \subset \cdots$. Then $\mathfrak{R} = \lim \mathfrak{R}_\nu = \mathfrak{R}_1 \cup \mathfrak{R}_2 \cup \mathfrak{R}_3 \cup \cdots$ is an open d. o. i.

We wish to show that $\mathfrak{R}$ *is compact.* This will follow from the following definition and lemma.

Definition. An element $a \in \mathfrak{F}$ is *integral* if its power sequence $\{a^\nu\}$ has no divergent subsequence.

---

[19] For a proof for hypercomplex systems with a finite basis see L. E. Dickson, *Algebras and their arithmetics*, Chicago, p. 47, or German edition, p. 97.

[20] $\mathfrak{P}^\nu$ is the smallest ring containing all the products $b_1 b_2 \cdots b_\nu$ where $b_\mu \in \mathfrak{P}$.

[21] J. H. M. Wedderburn, "A theorem on finite algebras," *Transactions of the American Mathematical Society*, vol. 6 (1905), pp. 349-352.

LEMMA 8.1. *A subring $\mathfrak{R}$ of $\mathfrak{F}$ containing only integral elements is compact.*

If $\{a_\nu\} \subset \mathfrak{R}$ and $a_\nu \to \infty$ then $a_\nu^{-1} \to 0$. Hence for a sufficiently large N, $a_N^{-1} \in \mathfrak{P}_1$ and so $\lim_{\mu\to\infty} a_N^{-\mu} = 0$. But this implies that $\lim_{\mu\to\infty} a_N^{\mu} = \infty$ which is impossible.

Since $\mathfrak{R} = \lim \mathfrak{R}_\nu$ as well as the $\mathfrak{R}_\nu$ contains integral elements only, it is compact. Since $\mathfrak{R}$ is c. o., it follows by the argument of Theorem 7.1 that $\mathfrak{R} = \mathfrak{R}_N$ for N sufficiently large and hence $\mathfrak{R}(\mathfrak{P}) = \mathfrak{R}$ where $\mathfrak{P}$ is the prime ideal of $\mathfrak{R}$.

Let $\mathfrak{P}^*$ denote the set of inverses of the elements of $\mathfrak{P}$.

THEOREM 8.1. $\mathfrak{F} = \mathfrak{R} \cup \mathfrak{P}^*$.

If $b \in \mathfrak{F}|\mathfrak{R} = \mathfrak{F}|\mathfrak{R}(\mathfrak{P})$, then there exists an element c_1 of $\mathfrak{P}$ such that $b_1 = bc_1 \in \mathfrak{F}|\mathfrak{P}$. If $b_1 \in \mathfrak{F}|\mathfrak{R}$, we may repeat this process and obtain a c_2 in $\mathfrak{P}$ such that $b_1c_2 = bc_1c_2 \in \mathfrak{F}|\mathfrak{P}$ and so on. Since $\mathfrak{P}^\nu \to 0$ the sequence $c_1, c_1c_2, c_1c_2c_3, \cdots$, if infinite, converges to 0 contradicting $bc_1c_2 \cdots c_\nu \in \mathfrak{F}|\mathfrak{P}$. Thus the sequence of c's breaks off after, say, r steps and we obtain $bc_1c_2 \cdots c_r = u \in \mathfrak{R}|\mathfrak{P}$. Hence $b^{-1} = c_1c_2 \cdots c_ru^{-1} \in \mathfrak{P}$, since $u^{-1} \in \mathfrak{R}|\mathfrak{P}$. Thus $\mathfrak{F}|\mathfrak{R} = \mathfrak{P}^*$.

This theorem shows that $\mathfrak{R}$ may be characterized topologically as the totality of integral elements of $\mathfrak{F}$. We have then

THEOREM 8.2. *$\mathfrak{R}$ is the only maximal c. o. d. o. i. in $\mathfrak{F}$.*

If $f \in \mathfrak{P}^\nu|\mathfrak{P}^{\nu+1}$, we say that f has *exponent* ν and if $f^{-1} \in \mathfrak{P}^\nu|\mathfrak{P}^{\nu+1}$, f has *exponent* $-\nu$. Define the value $|f| = \gamma^{\exp f}$ where $f \neq 0$ and $0 < \gamma < 1$ and $|0| = 0$.

THEOREM 8.3. *$|f|$ gives a non-archimedean valuation (Bewertung) of $\mathfrak{F}$: $|f+g| \leqq \max(|f|, |g|)$, $|fg| = |f| \cdot |g|$.*

We choose an element x in $\mathfrak{P}|\mathfrak{P}^2$ and we shall show that if f has exponent ν then $f = f_0x^\nu$ where $f_0 \in \mathfrak{R}|\mathfrak{P}$. First if f and g have exponent ν, fg^{-1} and $g^{-1}f$ have exponent 0. It is sufficient to prove this for $\nu > 0$. Now $fg^{-1} \in \mathfrak{P}$ implies $f \in \mathfrak{P}g \subset \mathfrak{P}^{\nu+1}$ which is contrary to $\exp f = \nu$. Similarly $fg^{-1} \in \mathfrak{P}^*$ is impossible and hence $fg^{-1} \in \mathfrak{R}|\mathfrak{P}$ and $\exp fg^{-1} = 0$. For the same reasons $\exp g^{-1}f = 0$. We observe next that $\exp x^\nu = \nu$. For if $\nu > 0$ $\exp x^\nu \geqq \nu$ and if $> \nu$ then $x^\nu = \Sigma y_1y_2 \cdots y_\mu$ where $y \in \mathfrak{P}$ and $\mu > \nu$. Then $1 = \Sigma x^{-\nu}y_1y_2 \cdots y_\mu$ and since $x^{-1}y \in \mathfrak{R}$ this implies that $1 \in \mathfrak{P}$ which is impossible. Thus $\exp(fx^{-\nu} = f_0) = 0$ and $f = f_0x^\nu$, $f_0 \in \mathfrak{R}|\mathfrak{P}$.

Suppose $f = f_0 x^\nu$, $g = g_0 x^\mu$ where $f_0, g_0 \in \mathfrak{R}|\mathfrak{P}$ and $\nu \leqq \mu$. Then

$$f + g = (f_0 + g_0 x^{\mu-\nu}) x^\nu \qquad (f_0 + g_0 x^{\mu-\nu}) \in \mathfrak{R}.$$

Hence

$$\exp(f+g) \geqq \min(\exp f, \exp g) \qquad |f+g| \leqq \max(|f|, |g|).$$

$$fg = (f_0 x^\nu g_0 x^{-\nu}) x^{\nu+\mu} \qquad (f_0 x^\nu g_0 x^{-\nu}) \in \mathfrak{R}|\mathfrak{P}$$

and so

$$\exp(fg) = \exp f + \exp g \qquad |fg| = |f| \cdot |g|.$$

$\mathfrak{J}$ is a *fractional left-(right-) ideal in* $\mathfrak{F}$ if it is a subgroup and $\mathfrak{R}\mathfrak{J} \subset \mathfrak{J}$ ($\mathfrak{J}\mathfrak{R} \subset \mathfrak{J}$). It follows easily that $\mathfrak{J}$ is open. The inverse $\mathfrak{J}^{-1}$ of $\mathfrak{J}$ is the set of elements a such that $a\mathfrak{J} \subset \mathfrak{R}$, or $\mathfrak{J}^{-1}\mathfrak{J} \subset \mathfrak{R}$. $\mathfrak{J}^{-1}$ is evidently a fractional left-ideal. If $\mathfrak{J} \neq (0), \neq 1$ then $\mathfrak{J}$ is called *proper*.

THEOREM 8.4. *Every proper fractional left-(right-) ideal is two-sided, principal and a power of* $\mathfrak{P}$.

If $b \in \mathfrak{J}$ and c has exponent $\geqq \exp b$, then $\exp cb^{-1} \geqq 0$, and $cb^{-1} \in \mathfrak{R}$. Hence $c \in \mathfrak{R}b \subset \mathfrak{J}$. If $\mathfrak{J}$ contains elements of arbitrarily small exponent, $\mathfrak{J} = \mathfrak{F}$. Otherwise let b have the smallest exponent of the elements of $\mathfrak{J}$. It follows that $\mathfrak{J}$ is the set of elements of $\mathfrak{F}$ of exponent $\geqq \exp b = k$ and hence $\mathfrak{J} = \mathfrak{P}^k = (b)$.

It follows directly from this theorem that $\mathfrak{J}\mathfrak{J}^{-1} = \mathfrak{R}$ as well as $\mathfrak{J}^{-1}\mathfrak{J}$. If $\mathfrak{F}$ has characteristic 0, $p\mathfrak{R}$ is an ideal $\neq 0$ and hence is $= \mathfrak{P}^e$ where $e \geqq 1$.

9. In this final section we shall apply the arithmetic results obtained in § 8 to obtain the structure of $\mathfrak{F}$.

Let p^n be the order of the finite field $\mathfrak{R} - \mathfrak{P}$ of characteristic p. $\mathfrak{R} - \mathfrak{P}$ contains a primitive q-th root of unity where $q = p^n - 1$, i. e., there exists a $u_0 \in \mathfrak{R}$ such that $u_0^q \equiv 1 (\mathfrak{P})$ and q is the smallest power for which such a congruence holds. Let $x \equiv 0 (\mathfrak{P})$ but $\not\equiv 0 (\mathfrak{P}^2)$. Then the correspondence $a \leftrightarrow xax^{-1}$ where $a \in \mathfrak{R}$ determines an automorphism of the Galois group of $\mathfrak{R} - \mathfrak{P}$. It follows that

$$xu_0x^{-1} \equiv u_0^s (\mathfrak{P}), \quad \text{where} \quad s = p^t.$$

THEOREM 9.1. $\mathfrak{F}$ *contains a primitive q-th root of unity.*

We shall determine $u \in \mathfrak{R}|\mathfrak{P}$ such that $u^q = 1$ and $u \equiv u_0 (\mathfrak{P})$. Since u_0 is a primitive q-th root of unity mod $\mathfrak{P}$ it will follow that u is primitive. Let r be the exponent of s mod q, i. e., r is the least positive integer for which $s^r \equiv 1 (q)$. Then r is also the least positive integer satisfying

$$(**) \qquad x^r u_0 x^{-r} \equiv u_0(\mathfrak{P}).$$

If $u_0^q \not\equiv 1$, suppose $u_0^q \equiv 1(\mathfrak{P}^k)$, $\not\equiv 1(\mathfrak{P}^{k+1})$. Then $u_0^q - 1 = v_0 x^k$ where $v_0 \in \mathfrak{R}|\mathfrak{P}$. Since $v_0 u_0 \equiv u_0 v_0(\mathfrak{P})$ and $(u_0^q - 1)u_0 = u_0(u_0^q - 1)$, we have $x^{-k}u_0 x^k \equiv u_0(\mathfrak{P})$ and hence by (**) $k \equiv 0(r)$, say $k = rl$. Thus

$$u_0^q \equiv 1 + wx^{rl}(\mathfrak{P}^{rl+1}) \qquad w \not\equiv 0(\mathfrak{P}).$$

Set $u_1 = u_0 + yx^{rl}$ where y is to be determined $\in \mathfrak{R}$ so that $u_1^q \equiv 1(\mathfrak{P}^{rl_1})$ for $l_1 > l$. This requires in particular that

$$(*^*_*) \qquad u_1^q \equiv 1 + (w + qu_0^{q-1}y)x^{rl} \equiv 1(\mathfrak{P}^{rl+1})$$

by (**) and the commutativity of u_0 and y mod $\mathfrak{P}$. Since q and $u_0^{q-1} \not\equiv 0(\mathfrak{P})$, y may be chosen so that $(w + qu_0^{q-1}y) \equiv 0(\mathfrak{P})$ and will then satisfy $(*^*_*)$. But then we will necessarily have $u_1^q \equiv 1(\mathfrak{P}^{rl_1})$ for $l_1 > l$. If u_1 is not a q-th root of one we may repeat this process with it in place of u_0. In this way we either obtain a q-th root of one after a finite number of steps or else an infinite sequence $\{u_\nu\}$ such that

$$u_\nu \equiv u_{\nu-1}(\mathfrak{P}^{rl_{\nu-1}}) \qquad u_\nu^q \equiv 1\,(\mathfrak{P}^{rl_\nu})$$

where $l < l_1 < l_2 < \cdots$. The $\{u_\nu\}$ converge to a limit u having the desired properties:

$$u^q = 1 \qquad u \equiv u_\nu(\mathfrak{P}^{rl_{\nu-1}}) \qquad \dot{x}ux^{-1} \equiv u^s(\mathfrak{P}).$$

THEOREM 9.2. *x may be normalized so that $xux^{-1} = u^s$, $x \in \mathfrak{P}|\mathfrak{P}^2$.*

This theorem has been given by Hasse for p-adic fields. Moreover his proof (H. p. 511) goes over word for word for the present more general case.

THEOREM 9.3. *Every element of $\mathfrak{F}$ may be represented uniquely in the form $v = (v_0 + v_1 x + v_2 x^2 + \cdots)x^{-k}$ where $k \geqq 0$ and $v_\nu = u^{k_\nu}$ or 0.*

If $v \equiv 0(\mathfrak{P}^{-k})$ for $k \geqq 0$, then $vx^k \equiv 0(\mathfrak{R})$ and hence $vx^k \equiv v_0(\mathfrak{P})$ where $v_0 = u^{k_0}$ or 0. Now $(vx^k - v_0)x^{-1} \equiv v_1(\mathfrak{P})$ where $v_1 = u^{k_1}$ or 0 and so on. We obtain thus a sequence $v_\nu = u^{k_\nu}$ or 0 such that

$$vx^k \equiv v_0 + v_1 x + v_2 x^2 + \cdots + v_\nu x^\nu(\mathfrak{P}^{\nu+1})$$

and hence

$$v = (v_0 + v_1 x + v_2 x^2 + \cdots)x^{-k}.$$

If $(\sum_{\nu=0}^{\infty} v_\nu x^\nu)x^{-k} = (\sum_{\nu=0}^{\infty} v'_\nu x^\nu)x^{-l}$, evidently $k = l$ so that $\Sigma v_\nu x^\nu = \Sigma v'_\nu x^\nu$. Hence

$v_0 \equiv v'_0(\mathfrak{P})$ and so $v_0 = v'_0$ since v_0 and v'_0 are members of the complete set of incongruent residues $0, u, u^2, \cdots, u^q \bmod \mathfrak{P}$. Likewise $v_1 = v'_1, v_2 = v'_2, \cdots$.

THEOREM 9.4. *The centrum $\mathfrak{C}$ of $\mathfrak{F}$ consists of all elements of the form $(c_0 + c_1x^r + c_2x^{2r} + \cdots)x^{-kr}$ where $c_\nu = 0$ or $u^{k\nu}$ and commutes with x.*

If $c = \sum_{\nu=-k}^{\infty} c_\nu x^\nu$, $u^{-1}cu = c$ implies $c_\nu = c_\nu u^{s^\nu - 1}$. Hence if $c_\nu \neq 0$, $u^{s^\nu - 1} = 1$ and $s^\nu \equiv 1(q)$ or $\nu \equiv 0(r)$. Further $xcx^{-1} = c$ implies that $c_\nu{}^s = c_\nu$ and hence c_ν commutes with x. Thus the two conditions that c commutes with x and with u imply that c has the form $\sum_{\nu=-k}^{\infty} c_\nu x^{\nu r}$, $c_\nu x = xc_\nu$. Since these two conditions insure that c belongs to the centrum $\mathfrak{C}$, we have the theorem.

The above proof shows also that the elements of $\mathfrak{F}$ commutative with u are all of the form $\sum_{\nu=-k}^{\infty} v_\nu x^{r\nu}$ where $v_\nu = u^{k\nu}$ or 0. These elements form a commutative subfield $\mathfrak{C}' \supset \mathfrak{C}$. Since $\mathfrak{C}' = \mathfrak{C}(u)$ it has a finite basis over $\mathfrak{C}$.

THEOREM 9.5. *$\mathfrak{C}'$ is a cyclic field of degree r over $\mathfrak{C}$.*

The automorphism $a \leftrightarrow xax^{-1}$ of $\mathfrak{F}$ leaves $\mathfrak{C}'$ invariant (though not element-wise) and hence induces an automorphism S in $\mathfrak{C}'$. The elements of $\mathfrak{C}'$ invariant under S are precisely the elements commutative with x and hence belonging to $\mathfrak{C}$. Since x^r is the smallest power of x commutative with u, the order of S is r. Hence $\mathfrak{C}'$ is cyclic of degree r over $\mathfrak{C}$ with S as generating automorphism of its Galois group relative to $\mathfrak{C}$.

THEOREM 9.6. *$\mathfrak{F}$ is a cyclic algebra over its centrum.*

For, by Theorem 9.5, u satisfies an irreducible equation of degree r, $\phi_r(u) = 0$ having coefficients in $\mathfrak{C}$. Thus

$$\phi_r(u) = 0 \qquad xu = u^s x \qquad (s = p^t) \qquad x^r = z \in \mathfrak{C}$$

gives a description of the algebra $\mathfrak{F}$ relative to its centrum $\mathfrak{C}$.

To complete the description of $\mathfrak{F}$ it is necessary to give the structure of $\mathfrak{C}$. Since $\mathfrak{C}$ is closed it is a l. c. s. t. d. commutative field and hence as mentioned in the introduction its form has been described by v. Dantzig and by Hasse and Schmidt. We shall merely state the results here and sketch briefly their derivation.

For the case $\chi(\mathfrak{F}) = 0$ we have seen that $\mathfrak{F}$ and hence also $\mathfrak{C}$ has finite order over the p-adic field P_p. From considerations analogous to those of

Theorems 9.1 and 9.2 it follows that $\mathfrak{C}$ is generated by a $p^m - 1$ $(m \geqq 1)$ root of one and a second element z $(= x^r)$. Certain additional normalizations may be made. The reader is referred to Hasse's paper (H. p. 514) for these.

If $\chi(\mathfrak{F}) = p$, we have

THEOREM 9.7. *The powers of u and 0 form a finite field $\mathbf{K}_{p^n}$ of p^n elements.*

It suffices to show that $v \equiv u^j + u^k = u^l$ or 0 and $-u^k = u^m$. Suppose $v \neq 0$. Since v is algebraic mod p (satisfies an algebraic equation whose coefficients are in the field $\mathbf{K}_p$ of residues mod p), it is a root of unity and hence $\epsilon\, \mathfrak{R}|\mathfrak{P}$. If $v \equiv u^l \pmod{\mathfrak{P}}$, $v - u^l$ is algebraic mod p and $\equiv 0 \pmod{\mathfrak{P}}$. This is impossible unless $v - u^l = 0$. Similarly we may show that $-u^k$ is $=$ some u^m.

By the same argument we see that x is transcendental mod p. Thus $\mathfrak{F}$ is uniquely determined by the field $\mathbf{K}_{p^n}$ and its automorphism $a \leftrightarrow a^s$ $(s = p^t)$. For when these are given we have $\mathfrak{F}$ determined as the set of integral power series in x with a finite number of negative powers and coefficients in $\mathbf{K}_{p^n}$ where the multiplication is determined by $xa = a^s x (a \epsilon \mathbf{K}_{p^n})$.

PRINCETON UNIVERSITY.

ON SINGULAR FOURIER-STIELTJES TRANSFORMS.

By RICHARD KERSHNER.

Let $\sigma(x)$, where $-\infty < x < +\infty$, be continuous, bounded and monotone non-decreasing and not a constant. If $\mu(S)$ denotes the measure of the set S of points x at which $\sigma(x)$ is actually increasing then $\mu(S) > 0$ is a necessary condition for the absolute continuity of $\sigma(x)$. If $\sigma(x)$ is absolutely continuous, then

$$(1) \qquad L(t,\sigma) \to 0 \text{ as } t \to \pm\infty, \text{ where } L(t,\sigma) = \int_{-\infty}^{\infty} e^{itx} d\sigma(x).$$

The question, under what additional condition does (1) imply the absolute continuity of $\sigma(x)$, is related to the problem of sets of uniqueness in the theory of trigonometrical series;[1] hence it is to be expected that the absolute continuity of σ must depend on arithmetical details. The present note furnishes among other things, an illustration to this situation.

It is known[2] that there exists for every positive $a < 1$ a continuous monotone $\sigma(x) = \sigma_a(x)$ such that

$$(2) \qquad L(t,\sigma_a) = \prod_{n=1}^{\infty} \cos(a^n t), \qquad -\infty < t < +\infty,$$

and that[3] the measure $\mu(S)$ of the corresponding set $S = S_a$ is zero whenever $0 < a < 1/2$. Now it will be shown that if a is a rational number in the latter interval, then (1) is satisfied by $\sigma = \sigma_a$ if and only if a^{-1} is not an integer. More precisely

$$(3) \quad L(t,\sigma_a) = O[(\log|t|)^{-\gamma}], \quad \gamma = \frac{-\log\cos(\pi/2q)}{\log(2\log q/\log p)} > 0, \; t \to \pm\infty$$

if $0 < a = p/q < 1/2$, where p and q are relatively prime and $p > 1$. In particular *there exist continuous monotone bounded functions σ which satisfy* (1) *and* $\mu(S) = 0$ *and are not absolutely continuous.* This implies a result of Menchoff[4] who considered the corresponding question for the case of the Fourier-Stieltjes coefficients where $t \to \pm\infty$ through integral values, rather than continuously.

---

[1] Cf. A. Zygmund, *Trigonometrical Series* (1935), p. 291 et seq.

[2] Cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[3] Cf. R. Kershner and A. Wintner, "On symmetric Bernoulli convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 541-548.

[4] Cf. A. Zygmund, *ibid.*

If $a = 1/3$, then σ_a is the Cantor function for which it is known that (1) does not hold.[5] That (1) does not hold for $\sigma = \sigma_a$ $(a = 1/4, 1/5, \cdots)$, can be shown in exactly the same way as is usually done for the Cantor function.[6] Hence it may be assumed that $p > 1$. In the proof of (3) it will not be necessary to assume that $0 < a < 1/2$ but only $0 < a < 1$, i. e. $q > p$. It is clearly sufficient to consider $t > 0$ since $L(t, \sigma_a)$ is an even function.

Let λ denote the set of all points $t > 0$ which are within a distance $\pi/2q$ of an integral multiple of π so that

$$| \cos t | > \cos \pi/2q$$

if and only if t is in λ. Let $t > \pi$ and let $h_0 = h_0(t) \geqq 0$ be the unique integer for which

$$q^{h_0+1} > t/\pi \geqq q^{h_0}. \tag{4}$$

Now either t is not in λ or there is a unique integer $k = k(t)$ such that

$$t = k\pi + \delta, \qquad | \delta | < \pi/2q \tag{5}$$

where $k < q^{h_0+2}$ by definition of h_0. In the latter case write

$$k = q^J l \tag{6}$$

where $J = J(t)$ is the exponent of the highest power of q contained in k, so that $0 \leqq J \leqq h_0 + 1$ and q does not divide l. By (5) and (6) we have

$$t(p/q)^{J+1} = \pi p^{J+1} l/q + \delta(p/q)^{J+1}.$$

Since p and q are relatively prime and q does not divide l then $p^{J+1}l/q$ must differ from an integer by at least $1/q$. And since

$$| \delta(p/q)^{J+1} | < | \delta | < \pi/2q$$

we see that $t(p/q)^{J+1}$ differs from an integral multiple of π by at least $\pi/q - \pi/2q = \pi/2q$ and hence that $t(p/q)^{J+1}$ is not in λ. Thus for any t satisfying (4) we know that $t(p/q)^{m_0}$ is not in λ for some m_0 such that $0 \leqq m_0 \leqq h_0 + 1$. Furthermore from (4) we get

$$p^{h_0+2}/q > t(p/q)^{h_0+2}/\pi \geqq p^{h_0+2}/q^2. \tag{7}$$

Now let t be chosen so large, for a fixed $n > 2$, that $h_0 = h_0(t)$ satisfies

$$h_0 + 2 \geqq (2 \log q/\log p)^n$$

---

[5] Cf. B. Jessen and A. Wintner, *ibid.*, where further references are given.

[6] Note that this proof breaks down in the case $a = 1/2$ since $\cos \pi/2 = 0$ and in fact in this case (1) holds. Cf. B. Jessen and A. Wintner, *ibid.*

this choice of t being possible in view of (4). Then

$$(h_0+2)(\log p/\log q) \geqq 2(2\log q/\log p)^{n-1} > (2\log q/\log p)^{n-1}+1.$$

Hence we can find an integer $g = g(t) \geqq 0$ such that

$$(h_0+2)(\log p/\log q) \geqq g+2 \geqq (2\log q/\log p)^{n-1}.$$

The first of these inequalities may be written

$$p^{h_0+2}/q^2 \geqq q^g$$

and comparing this with (7) we see that there is a unique integer $h_1 = h_1(t)$ such that

$$q^{h_1+1} > t(p/q)^{h_0+2}/\pi \geqq q^{h_1}$$

and

$$(h_1+2) \geqq (g+2) \geqq (2\log q/\log p)^{n-1}.$$

Applying to $t(p/q)^{h_0+2}$ the argument that we applied above to t we see that $t(p/q)^{m_1}$ is not in λ for some m_1 such that $h_0+2 \leqq m_1 \leqq h_0+h_1+3$ and that

$$q^{h_2+1} \geqq t(p/q)^{h_0+h_1+4}/\pi \geqq q^{h_2}$$

where

$$h_2+2 \geqq (2\log q/\log p)^{n-2}.$$

Applying the above argument $n-1$ times it follows that if

$$t/\pi > q^h \quad \text{where} \quad h > (2\log q/\log p)^n \tag{8}$$

then there are at least $n-1$ distinct values of m such that

$$0 \leqq m \leqq h_0+h_1+\cdots+h_{n-1}+2^{n-1}-1$$

and that $t(p/q)^m$ is not in λ and consequently

$$|L(t,\sigma_a)| < \cos^{n-1}(\pi/2q). \tag{9}$$

Let

$$t/\pi \geqq q^{A^{n+1}} \quad \text{where} \quad A = A_{pq} = (2\log q/\log p) > 2. \tag{10}$$

Then certainly there is an h satisfying (8).

Let $n(t)$ be the largest integer satisfying (10). Then

$$n(t) > \{\log[\log t - \log\pi] - \log\log q - 2\log A\}/\log A$$

and substitution into (9) gives the required appraisal (3).

THE JOHNS HOPKINS UNIVERSITY.